

CC composition for detecting hydrogen peroxide, e.g. for diagnosis of
 CC uraemia, chronic nephritis, giantism and tonic muscular dystrophy. The
 CC enzyme has a lower Km value than prior art creatine amidinohydrolase (cf.
 CC US 5451520). (Updated on 17-OCT-2003 to standardise OS field)

XX Sequence 1212 BP; 244 A; 400 C; 358 G; 210 T; 0 U; 0 Other;

Query Match 99.9%; Score 1210.4; DB 2; Length 1212;
 Best Local Similarity 99.9%; Pred. No. 4.2e-196;
 Matches 1211; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 ATGACTGACGACATCTTTCACGATGAAATGGCAACACGGCGAAGATTATTTGCGG 60
DB 1 ATGACTGACGACATCTTTCACGATGAAATGGCAACACGGCGAAGATTATTTGCGG 60
QY 61 TTTTGGATGCGGAGATGACCGCGCGCAAAAGAGGTTGCGGGCTGATGGCGAAGAC 120
DB 61 TTTTGGATGCGGAGATGACCGCGCGCAAAAGAGGTTGCGGGCTGATGGCGAAGAC 120
QY 121 AATGTGATGCGGCGCTGTTCACCTCTTATCACTGATCACTACTATTCGGCTGGCTG 180
DB 121 AATGTGATGCGGCGCTGTTCACCTCTTATCACTGATCACTACTATTCGGCTGGCTG 180
QY 181 TACTGCTATTTGGACGCAAGTACGGCATGCTCATGACCAACAACGCCACGACGATT 240
DB 181 TACTGCTATTTGGACGCAAGTACGGCATGCTCATGACCAACAACGCCACGACGATT 240
QY 241 TGGGCGGATGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
DB 241 TGGGCGGATGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
QY 301 ACCGACTGCGCGCGCGCAATTTCTATCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
DB 301 ACCGACTGCGCGCGCGCGCAATTTCTATCGCGCGCGCGCGCGCGCGCGCGCGCG 360
QY 361 CGCATGCGGATGAGTTGCAACACGTCATCTTCGACCTTCGCGCGCGCGCGCGCG 420
DB 361 CGCATGCGGATGAGTTGCAACACGTCATCTTCGACCTTCGCGCGCGCGCGCGCG 420
QY 421 CTACCGGGGGTGAATTCGTCACATCAACCGCGCGCGCGCGCGCGCGCGCGCGCG 480
DB 421 CTACCGGGGGTGAATTCGTCACATCAACCGCGCGCGCGCGCGCGCGCGCGCGCG 480
QY 481 TGGCTCGAAGACAGAGCTGATCCGCGAAGGCGCGCGCGCGCGCGCGCGCGCGCG 540
DB 481 TGGCTCGAAGACAGAGCTGATCCGCGAAGGCGCGCGCGCGCGCGCGCGCGCGCG 540
QY 541 GCGTCGCGCGCTGCATCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
DB 541 GCGTCGCGCGCTGCATCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
QY 601 AATGCGATGATCCGCGAGATCCCAATGCTTCCCTTCGTGAGCTGATGACACTGG 660
DB 601 AATGCGATGATCCGCGAGATCCCAATGCTTCCCTTCGTGAGCTGATGACACTGG 660
QY 661 ACCGTGTTCCATCCGCGCATCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
DB 661 ACCGTGTTCCATCCGCGCATCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
QY 721 GTGCAATCCGCGCATCTTTCGCTCAACACCTTCCGATGATCTTCCGCTACTACAC 780
DB 721 GTGCAATCCGCGCATCTTTCGCTCAACACCTTCCGATGATCTTCCGCTACTACAC 780
QY 781 GCGCTGAGAGCGACGCTGTTTCGCAACATGTCATGATGACGCGCGCGCGCGCGCG 840
DB 781 GCGCTGAGAGCGACGCTGTTTCGCAACATGTCATGATGACGCGCGCGCGCGCGCG 840
QY 841 AAGAACGTGCGCTGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
DB 841 AAGAACGTGCGCTGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
QY 901 GACATGCGCATGACGCTCAACAGATGTATCCGCGAGTGGACCTGCTGAATACGCTCC 960
DB 901 GACATGCGCATGACGCTCAACAGATGTATCCGCGAGTGGACCTGCTGAATACGCTCC 960

```

```

DB 901 GACATGCGCATGACGCTCAACAGATGTATCCGCGAGTGGACCTGCTGAATACGCTCC 960
QY 961 TTGGCTATGACCACTCTTCGCGCGCTGTCGCACTACTACGTCGCGAGCGCGCTG 1020
DB 961 TTGGCTATGACCACTCTTCGCGCGCTGTCGCACTACTACGTCGCGAGCGCGCTG 1020
QY 1021 GAGCTGCGGAGACATGACACCGAGCTGAAGCCCGCGCATGGTGTCTTCATGAGCCG 1080
DB 1021 GAGCTGCGGAGACATGACACCGAGCTGAAGCCCGCGCATGGTGTCTTCATGAGCCG 1080
QY 1081 ATGCTGATGCTCCCGAGAGGCAATGCCCGGTGCGCGCGCTATGCGACGACATCTG 1140
DB 1081 ATGCTGATGCTCCCGAGAGGCAATGCCCGGTGCGCGCGCTATGCGACGACATCTG 1140
QY 1141 ATGCTGAGGAGAGACGCTCCGAGAACATCACCGGCTTCCGTTGCGTCCGGAACAAC 1200
DB 1141 ATGCTGAGGAGAGACGCTCCGAGAACATCACCGGCTTCCGTTGCGTCCGGAACAAC 1200
QY 1201 ATCATCCGCAAC 1212
DB 1201 ATCATCCGCAAC 1212

RESULT 3
AAV35699
ID AAV35699 standard; DNA, 1212 BP.
XX
AC AAV35699;
XX
DT 16-SEP-1998 (first entry)
XX
DE Stable creatine amidinohydrazase encoding DNA.
XX
KW Creatine amidinohydrazase; mutant; stable; enzyme; diagnostic agent; ds.
XX
OS Alcaligenes faecalis.
XX
FH Key Location/Qualifiers
FT CDS 1..1212
FT /tag=a
FT /transl_except=(pos:433..435, aa:Glu)
FT /product="Stable creatine amidinohydrazase"
FT /note="the stop codon is not indicated"
XX
XX JP10174585-A.
XX
XX PD 30-JUN-1998.
XX
XX PF 17-DEC-1996; 96JP-00337027.
XX
XX PR 17-DEC-1996; 96JP-00337027.
XX
XX PA (TOYO) TOYOBO KK.
XX
XX DR MPI; 1998-421167/36.
XX
XX PT P-PSDB; AAM61905.
XX
XX PT New creatine amidino-hydrazase used as diagnostic agent - is more stable in
XX neutral buffer than wild type creatine amidino-hydrazase.
XX
XX PS Disclosure; Page 11-13; 14pp; Japanese.
XX
XX CC This DNA encodes a stable creatine amidinohydrazase which is a mutant
XX creatine amidinohydrazase and has improved long-term stability in a neutral
XX buffer compared to wild type creatine amidinohydrazase. A recombinant
XX plasmid containing the stable creatine amidinohydrazase gene can be used to
XX transform a cell for the recombinant production of the enzyme. This
XX stable creatine amidinohydrazase is useful as a diagnostic agent can be
XX produced commercially.
XX
SQ Sequence 1212 BP; 244 A; 400 C; 358 G; 210 T; 0 U; 0 Other;
Query Match 99.9%; Score 1210.4; DB 2; Length 1212;

```

Best Local Similarity 99.9%; Pred. No. 4,2e-196;
Matches 1211; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

OY 1 ATGACTGACGACATGTTGACGTGATGAATAATGACACACGCGAGAGAAATATTTCGCCG 60
DB 1 ATGACTGACGACATGTTGACGTGATGAATAATGACACACGCGAGAGAAATATTTCGCCG 60
OY 61 TTTTGGGATGCGGAGATGACCCGCCGCAAAAGAGAGTTCCGGGCTGATGGCCAAAGAC 120
DB 61 TTTTGGGATGCGGAGATGACCCGCCGCAAAAGAGAGTTCCGGGCTGATGGCCAAAGAC 120
OY 121 AATGTGATGCGGCGGTGTTCACTCTTATCACTGATCAACTATCTATTCGGGCTG 180
DB 121 AATGTGATGCGGCGGTGTTCACTCTTATCACTGATCAACTATCTATTCGGGCTG 180
OY 181 TACTGCTATTTGAGACGCAAGTACGCGATGATGACACACACACGCGACAGAT 240
DB 181 TACTGCTATTTGAGACGCAAGTACGCGATGATGACACACACACGCGACAGAT 240
OY 241 TCGGCGGCGATGACGCGCGCGACGCTTGGCGCGGACCTTGGGCGACACATCACTAC 300
DB 241 TCGGCGGCGATGACGCGCGCGACGCTTGGCGCGGACCTTGGGCGACACATCACTAC 300
OY 301 ACCGACTGGCGCGCGGACAAATTTCTATGCGGCGTGGCGCAGCTGACACAGGCGCGAAG 360
DB 301 ACCGACTGGCGCGCGGACAAATTTCTATGCGGCGTGGCGCAGCTGACACAGGCGCGAAG 360
OY 361 CGCATGCGCATGAGTTGACACACGTCATCTGACCTTCGCGCGCGACGCTGAGAGAGCC 420
DB 361 CGCATGCGCATGAGTTGACACACGTCATCTGACCTTCGCGCGCGACGCTGAGAGAGCC 420
OY 421 CTACCGGCGGTGAGTTGATGACATCAGCCAGCCCTGATGATGATGATGATGATGATGAT 480
DB 421 CTACCGGCGGTGAGTTGATGACATCAGCCAGCCCTGATGATGATGATGATGATGATGAT 480
OY 481 TCGGCTGAGAGACAGAGAGTTCGCGAGAGGCGCGGCTGATGATGATGATGATGATGATGAT 540
DB 481 TCGGCTGAGAGACAGAGAGTTCGCGAGAGGCGCGGCTGATGATGATGATGATGATGATGAT 540
OY 541 GCGTCGCGGCTGACATCAAGGCGCGGCTGCGCGAGACAGAGAGTGGCATTCGCGACACC 600
DB 541 GCGTCGCGGCTGACATCAAGGCGCGGCTGCGCGAGACAGAGAGTGGCATTCGCGACACC 600
OY 601 AATGCGATGATCGCGAGATTCGCAAAATGTTCCCTTCTGATGATGATGATGATGATGATGAT 660
DB 601 AATGCGATGATCGCGAGATTCGCAAAATGTTCCCTTCTGATGATGATGATGATGATGATGAT 660
OY 661 AACTGCTTCATTCGCGGATCAACACGAGCGGCGGCAATTCGCGTCAACACGCGCATC 720
DB 661 AACTGCTTCATTCGCGGATCAACACGAGCGGCGGCAATTCGCGTCAACACGCGCATC 720
OY 721 GTGCAATCCGCGGACATCTCTTCTGCTCAACACCTTCCGATGATCTTCCGCTACTACACC 780
DB 721 GTGCAATCCGCGGACATCTCTTCTGCTCAACACCTTCCGATGATCTTCCGCTACTACACC 780
OY 781 GCGCTGAGAGCGACGCTGTTCTGCGACATGTCGATGAGCGCGACGCTTCAACTCTGGGAG 840
DB 781 GCGCTGAGAGCGACGCTGTTCTGCGACATGTCGATGAGCGCGACGCTTCAACTCTGGGAG 840
OY 841 AAGAAAGTGGCGGTGATGCGCGGCGGCTGAGCTATCAAGCGCGGCGCGGCTGCAAG 900
DB 841 AAGAAAGTGGCGGTGATGCGCGGCGGCTGAGCTATCAAGCGCGGCGCGGCTGCAAG 900
OY 901 GACATGCGCATGAGCTCAACGAGATGTAACGCGAGTGGGACCTGCTGAAGTACCGCTCC 960
DB 901 GACATGCGCATGAGCTCAACGAGATGTAACGCGAGTGGGACCTGCTGAAGTACCGCTCC 960
OY 961 TTGCGGTATGCGCATCTCTTCCGCGCTGCTGCGACATCACTACGCTGCGAGGCGCGCTG 1020
DB 961 TTGCGGTATGCGCATCTCTTCCGCGCTGCTGCGACATCACTACGCTGCGAGGCGCGCTG 1020
OY 1021 GAGCTGCGGAGGACATCAACCGAGCTGAAGCGCGGAGTGGTCTTCATGAGCGG 1080
DB 1021 GAGCTGCGGAGGACATCAACCGAGCTGAAGCGCGGAGTGGTCTTCATGAGCGG 1080

```

```

DB 1021 GAGCTGCGGAGGACATCAACCGAGCTGAAGCGCGGAGTGGTCTTCATGAGCGG 1080
OY 1081 ATGGTGAATGCTCCCGGAGGAGCATGCCGCTTCGCGAGACAGACATCCG 1140
DB 1081 ATGGTGAATGCTCCCGGAGGAGCATGCCGCTTCGCGAGACAGACATCCG 1140
OY 1141 ATCGTGGGAGGAGACGCTGCGGAGAACATCAACGCGCTTCCGTTGCTGCGGAGACAAAC 1200
DB 1141 ATCGTGGGAGGAGACGCTGCGGAGAACATCAACGCGCTTCCGTTGCTGCGGAGACAAAC 1200
OY 1201 ATCATCCGCAAC 1212
DB 1201 ATCATCCGCAAC 1212

```

RESULT 4
AAS14742
ID AAS14742 standard; DNA; 1212 BP.
XX
AC AAS14742;
XX
DT 19-DEC-2001 (first entry)
XX
DS Creatine amidinohydrolase genomic DNA.
XX
KM Creatine amidinohydrolase; water; sarcosine; urea; creatinine; uraemia;
KM chronic nephritis; acute nephritis; tonic muscular dystrophy; giantism;
KM pigment absorbance; ds.
XX
OS Alcaligenes faecalis.
XX
FH Key Location/Qualifiers
FT 1..1212
FT CDS
FT /tag= a
FT /product= "A. faecalis creatine amidinohydrolase"
FT /transl_except= (pos:433..435, aa:Glu)
FT /partial
FT /note= "No stop codon"

EP132467-A2.
PD 12-SEP-2001.
XX
PF 13-FEB-1997; 2001EP-00113052.
XX
PR 13-FEB-1996; 96JP-00025435.
PR 13-FEB-1997; 97EP-00102270.
XX
PA (TOYM) TOYO BOSEKI KK.
XX
PI Sogabe A, Hattori T, Nishiya Y, Kawamura Y;
DR WPI; 2001-612481/71.
DR P-PSDB; AAU08727.
XX
PT New creatine amidinohydrolase, useful as a routine reagent for clinical
PT tests for determining creatine and creatinine in biological samples,
PT particularly useful in diagnosing diseases such as uremia or chronic
PT nephritis.
XX
PS Disclosure; Page 15-16; 21pp; English.
XX
CC The invention relates to Alcaligenes faecalis creatine amidinohydrolase,
CC which catalyses creatine and water to sarcosine and urea. Creatine
CC amidinohydrolase can be produced by culturing a microorganism producing
CC the protein in a nutrient medium and recovering the protein from the
CC resulting culture. Creatine amidinohydrolase is useful as a routine
CC reagent for clinical tests for determining creatine and creatinine in
CC biological samples. This is particularly useful in diagnosing diseases
CC such as uremia, chronic nephritis, acute nephritis, giantism and tonic
CC muscular dystrophy. The presence of creatine in a sample can be
CC determined by measuring an absorbance of a pigment produced by the
CC reaction of a reagent containing creatine amidinohydrolase with the

CC sample. This sequence represents genomic DNA encoding Alcaligenes
 CC faecalis creatine amidinohydrolase
 XX

Sequence 1212 BP; 244 A; 400 C; 358 G; 210 T; 0 U; 0 Other;

Query Match 99.9%; Score 1210.4; DB 4; Length 1212;

Best Local Similarity 99.9%; Pred. No. 4.2e-196;

Matches 1211; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 ATGACTGACGACATGTTGCGACGTGATGAATGGCAACAAGCGCGAAGAAATTATTTGCGCG 60
DB 1 ATGACTGACGACATGTTGCGACGTGATGAATGGCAACAAGCGCGAAGAAATTATTTGCGCG 60
QY 61 TTTTGGGATGCGGAGATGACCGCGCGCGCAAAAGAGCGTTGCGCGGTGAGTGGCAAGAAC 120
DB 61 TTTTGGGATGCGGAGATGACCGCGCGCGCAAAAGAGCGTTGCGCGGTGAGTGGCAAGAAC 120
QY 121 AATGTCGATGCGGCGCTGTTCACTCTTATCACTGATCAACTAATTCTCGGCTGCTG 180
DB 121 AATGTCGATGCGGCGCTGTTCACTCTTATCACTGATCAACTAATTCTCGGCTGCTG 180
QY 181 TACTGCTATTTCCGAGCGCAAGTACGGCATGTGTATGACCAACAACGCGACAGGAT 240
DB 181 TACTGCTATTTCCGAGCGCAAGTACGGCATGTGTATGACCAACAACGCGACAGGAT 240
QY 241 TCGGCGGCGGATCGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
DB 241 TCGGCGGCGGATCGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
QY 301 ACCGATGCGGCGCGCGCGCAATTTCTATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
DB 301 ACCGATGCGGCGCGCGCGCAATTTCTATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
QY 361 CGGATGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
DB 361 CGGATGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
QY 421 CTACCGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
DB 421 CTACCGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
QY 481 TCGCTCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
DB 481 TCGCTCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
QY 541 GCGTGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
DB 541 GCGTGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
QY 601 AATGCGATGATCGGCGAGATCGGCAATGCTCCCTGCGTGGAGCGTGAATGACACCTG 660
DB 601 AATGCGATGATCGGCGAGATCGGCAATGCTCCCTGCGTGGAGCGTGAATGACACCTG 660
QY 661 ACTGCGTTCAGTCCGCGCATCAACCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCAT 720
DB 661 ACTGCGTTCAGTCCGCGCATCAACCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCAT 720
QY 721 GTGCAATCCGCGCGCATCTTTGCTCAACACCTTCCCGATGATCTTCCGCTACTACAC 780
DB 721 GTGCAATCCGCGCGCATCTTTGCTCAACACCTTCCCGATGATCTTCCGCTACTACAC 780
QY 781 GGGCTGGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 840
DB 781 GGGCTGGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 840
QY 841 AAGAAAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
DB 841 AAGAAAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
QY 901 GACATGCGCATGAGCTCAACGAGATGTACCGCGGATGGGACCTGCTGAAGTACCGCTCC 960
DB 901 GACATGCGCATGAGCTCAACGAGATGTACCGCGGATGGGACCTGCTGAAGTACCGCTCC 960

```

```

QY 961 TTCGCTATGCGCATCTTTCGCGCGTGTGCGCACTACTACGTCGCGAGCGCGCGT 1020
DB 961 TTCGCTATGCGCATCTTTCGCGCGTGTGCGCACTACTACGTCGCGAGCGCGCGT 1020
QY 1021 GAGCTCGCGAGGACATCGACACCGAGCTGAAGCCCGCGCATGTGCTTCCATGAGCG 1080
DB 1021 GAGCTCGCGAGGACATCGACACCGAGCTGAAGCCCGCGCATGTGCTTCCATGAGCG 1080
QY 1081 ATGCTGATGCTCCCGAGGCGCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1140
DB 1081 ATGCTGATGCTCCCGAGGCGCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1140
QY 1141 ATGCTGCGGAGGACGCTCCCGAGACATCACCGCGCTTCCGCTCCGAGCAAC 1200
DB 1141 ATGCTGCGGAGGACGCTCCCGAGACATCACCGCGCTTCCGCTCCGAGCAAC 1200
QY 1201 ATCATCGCGAC 1212
DB 1201 ATCATCGCGAC 1212

```

RESULT 5

AA199856
 ID AA199856 standard; DNA; 1212 BP.

XX AA199856;

DT 07-AUG-2003 (revised)

DT 28-JUN-2002 (first entry)

XX Alkaligenes faecalis creatineamidinohydrolase encoding DNA.

KM Alkaligenes faecalis; TB3581; FERM P14237; creatineamidinohydrolase; ds.

XX Alcaligenes faecalis.

XX Key Location/Qualifiers

FT CDS 1..1212

FT FT /transl_except= (pos:433..435,aa:Glu)

FT FT /product= "creatineamidinohydrolase"

FT FT /note= "CDS lacks a stop codon"

PN JP2001252088-A.

XX 18-SEP-2001.

PF 16-MAY-1995; 2001JP-00051054.

PR 16-MAY-1995; 95JP-00117283.

PA (TOYO) TOYOBO KK.

XX WPI; 2002-003140/01.

DR P-PSDB; AAM51471.

XX A gene encoding creatineamidinohydrolase.

PT Claim 4; Page 10; 11pp; Japanese.

XX The invention relates to Alkaligenes faecalis TB3581 (FERM P-14237)

CC creatineamidinohydrolase and the encoding gene. The gene can be used for

CC the commercial preparation of creatineamidinohydrolase. (Updated on 07-

CC AUG-2003 to correct OS field.)

XX Sequence 1212 BP; 244 A; 400 C; 358 G; 210 T; 0 U; 0 Other;

QY Query Match 99.9%; Score 1210.4; DB 6; Length 1212;

Best Local Similarity 99.9%; Pred. No. 4.2e-196;

Matches 1211; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGACTGACGACATGTTGCGACGTGATGAATGGCAACAAGCGCGAAGAAATTATTTGCGCG 60

```

Db      1 ATGACGACGACATGTTGACGTTGATGAAATGGCAACAAGCGGAGAAAGATTATTCGCGG 60
Qy      61 TTTTGGATGCCGAGATGACCCCGCGCCAAAAGACGATTCGGGGCTGATGGCCAAAG 120
Db      61 TTTTGGATGCCGAGATGACCCCGCGCCAAAAGACGATTCGGGGCTGATGGCCAAAG 120
Qy      121 AATGCGATGGCGGCTGTTCACTTATCACTGATCACTAATTCGGGCTGCTG 180
Db      121 AATGCGATGGCGGCTGTTCACTTATCACTGATCACTAATTCGGGCTGCTG 180
Qy      181 TACTGCTATTTGCGACGCAAGTACGGCATGATCATGACCAACAACGCCACGAGATT 240
Db      181 TACTGCTATTTGCGACGCAAGTACGGCATGATCATGACCAACAACGCCACGAGATT 240
Qy      241 TCGGCGCGGATGACGCGCGCGCCAGCCCTGCGCGCGAGTTGGGCGACAACTACCTAC 300
Db      241 TCGGCGCGGATGACGCGCGCGCCAGCCCTGCGCGCGAGTTGGGCGACAACTACCTAC 300
Qy      301 ACCGACTGGCGCGCGGACAAATTTCTATGCGGCGGTGCGCGCAGCTGACCAAGGCGCAAG 360
Db      301 ACCGACTGGCGCGCGGACAAATTTCTATGCGGCGGTGCGCGCAGCTGACCAAGGCGCAAG 360
Qy      361 CGCATCGGATGACGAGTTGCAACCACTGCAATCTGCACTTCGCGCGCAGCTCGAGGAAAGCC 420
Db      361 CGCATCGGATGACGAGTTGCAACCACTGCAATCTGCACTTCGCGCGCAGCTCGAGGAAAGCC 420
Qy      421 CTACCGGCGGTGAGTTGCTGCAATCAAGCCAGCCCTGATGATGATGCGGCACTACAG 480
Db      421 CTACCGGCGGTGAGTTGCTGCAATCAAGCCAGCCCTGATGATGATGCGGCACTACAG 480
Qy      481 TCGCTCGAAGACAGAAAGCTGATTCGCGGAGGCGCGCGGTGATGATGATGCGGCGCGCG 540
Db      481 TCGCTCGAAGACAGAAAGCTGATTCGCGGAGGCGCGCGGTGATGATGATGCGGCGCGCG 540
Qy      541 GCGTCGCGGCTGCAATCAAGGCGCGCGGTGCGCGGAGCATGAAGTGGCGATCGCAACACC 600
Db      541 GCGTCGCGGCTGCAATCAAGGCGCGCGGTGCGCGGAGCATGAAGTGGCGATCGCAACACC 600
Qy      601 AATGCGATGATCGCGGAGATTCGCAATGTTCCCTTGTGAGAGTGAATGACACCTGAG 660
Db      601 AATGCGATGATCGCGGAGATTCGCAATGTTCCCTTGTGAGAGTGAATGACACCTGAG 660
Qy      661 ACCTGCTTCCATCGCGGATCAACACCGAGGCGGCAATCCGGTCAACAACGCGATC 720
Db      661 ACCTGCTTCCATCGCGGATCAACACCGAGGCGGCAATCCGGTCAACAACGCGATC 720
Qy      721 GTGCAATCGCGGACATCTTTGCTGCAACACCTTCCGATGATCTTCCGCTACTACACC 780
Db      721 GTGCAATCGCGGACATCTTTGCTGCAACACCTTCCGATGATCTTCCGCTACTACACC 780
Qy      781 GCGCTGAGACGACGCTGTTGTCGACATGTCGATGACGCCAGCTCGACATCTGGAG 840
Db      781 GCGCTGAGACGACGCTGTTGTCGACATGTCGATGACGCCAGCTCGACATCTGGAG 840
Qy      841 AAGAAGCTGCGGCTGATGCGCGCGGCTGAGCTGATCAAGCGCGGCGCGGCTGCGAG 900
Db      841 AAGAAGCTGCGGCTGATGCGCGCGGCTGAGCTGATCAAGCGCGGCGCGGCTGCGAG 900
Qy      901 GACATCGCATGACGCTCAACAGAGATGATCCGAGATGGAGACTGCTGAAGTACCGCTTC 960
Db      901 GACATCGCATGACGCTCAACAGAGATGATCCGAGATGGAGACTGCTGAAGTACCGCTTC 960
Qy      961 TTGCGCTATGGCACTCTTTCGCGGTGCTGTGCACTACTACGCTTGGAGAGCGCGCTG 1020
Db      961 TTGCGCTATGGCACTCTTTCGCGGTGCTGTGCACTACTACGCTTGGAGAGCGCGCTG 1020
Qy      1021 GAGCTGCGGAGAGACATGCAACCGAGCTGAAGCCCGGATGGTGTCTTCATGAGCCG 1080
Db      1021 GAGCTGCGGAGAGACATGCAACCGAGCTGAAGCCCGGATGGTGTCTTCATGAGCCG 1080
Qy      1081 ATGCTGATGCTGCGGAGGCAATGCGCGGTGCGGCGGCTATCGCAGACGACATCTCTG 1140

```

```

Db      1081 ATGCTGATGCTGCGGAGGCGGATGCCGCTCCGCGGCTATTCGACACGACATCTCTG 1140
Qy      1141 ATGCTGCGGAGAGACGCTGTCGAGAACATCAACCGGCTTCGCTTCGAGACACAAC 1200
Db      1141 ATGCTGCGGAGAGACGCTGTCGAGAACATCAACCGGCTTCGCTTCGAGACACAAC 1200
Qy      1201 ATCATCCGCAAC 1212
Db      1201 ATCATCCGCAAC 1212

RESULT 6
ABA93696
ID ABA93696 standard; DNA; 1212 BP.
XX
AC ABA93696;
XX
DT 29-APR-2002 (first entry)
XX
DE Creatine amidinohydrazase encoding DNA SEQ ID NO:2.
XX
KW Creatine amidinohydrazase; enzyme; mutant; stable; clinical diagnosis;
KW gene; ds.
XX
OS Alcaligenes faecalis.
XX
PH Key Location/Qualifiers
XX CDS 1..1212
XX FT /*tag= a
XX FT /product= "creatine amidinohydrazase"
XX FT /transl_except= (pos:433..435,aa:Glu)
XX FT /note= "no stop codon given"
XX
PN JP2001346594-A.
XX
PD 18-DEC-2001.
XX
PF 17-DEC-1996; 2001JP-00121708.
XX
PR 17-DEC-1996; 96JP-00337027.
XX
PA (TOYO) TOYOBO KK.
XX
DR WPI; 2002-145187/19.
XX P-PSDB; ABB05660.
XX
PT Mutant creatine amidinohydrazase for use in clinical diagnosis has long-
XX term stability in a neutral buffer solution.
XX
PS Claim 12; Page 12-13; 15pp; Japanese.
XX
CC The present invention describes a stable mutant creatine amidinohydrazase
CC having a long-term stability in a neutral buffer solution. Also described
CC are: (1) a creatine amidinohydrazase gene encoding the above stable
CC creatine amidinohydrazase; (2) a gene encoding a mutant creatine
CC amidinohydrazase having a long-term stability in a neutral buffer solution
CC compared to wild type creatine amidinohydrazase, in which at least one gene
CC among those found in a fully defined 1212 nucleotide sequence (the
CC present sequence) is replaced by another gene; (3) a recombinant plasmid
CC containing a gene encoding the above creatine amidinohydrazase; (4) a cell
CC transformed by the above plasmid; and (5) a method for the preparation of
CC a stable creatine amidinohydrazase in which the above cell is cultured in a
CC medium and creatine amidinohydrazase is collected. The creatine
CC amidinohydrazase is useful as a clinical diagnosing agent
XX
SQ Sequence 1212 BP; 244 A; 400 C; 358 G; 210 T; 0 U; 0 Other;
XX

Query Match 99.9%; Score 1210.4; DB 6; Length 1212;
Best Local Similarity 99.9%; Pred. No. 4.2e-196;
Matches 1211; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 ATGACTGACGACATTTGACGTTGATGAAATGGCAACAAGCGGAGAAAGATTATTCGCGG 60

```



```

QY 181 TACTGCTATTTCGAGCGCAAGTACGGCATGTGATGACCAACAACGCGACAGAGATT 240
DB 181 TACTGCTATTTCGAGCGCAAGTACGGCATGTGATGACCAACAACGCGACAGAGATT 240
QY 241 TCGGCGGCGCATCGACGGCGCGCAGCCCTGGCGCGCGAGCTTGGCGCAACATCACTTAC 300
DB 241 TCGGCGGCGCATCGACGGCGCGCAGCCCTGGCGCGCGAGCTTGGCGCAACATCACTTAC 300
QY 301 ACCGACTGCGCGCGCGCAATTTCTATTCGCGCGCTGCGCGAGCTGACCAACGCGCGCGCAAG 360
DB 301 ACCGACTGCGCGCGCGCAATTTCTATTCGCGCGCTGCGCGAGCTGACCAACGCGCGCGCAAG 360
QY 361 CGGATGCGGATCGAATTTGCAACGCTGCAACTTGGCGCGCGCGAGCTTGGCGCAAGCGC 420
DB 361 CGGATGCGGATCGAATTTGCAACGCTGCAACTTGGCGCGCGCGAGCTTGGCGCAAGCGC 420
QY 421 CTACCGGCGCTCGAGTTCTGCAATCAAGCGCGAGCTTGGCGAGTGGATGCGCAACATCAAG 480
DB 421 CTACCGGCGCTCGAGTTCTGCAATCAAGCGCGAGCTTGGCGAGTGGATGCGCAACATCAAG 480
QY 481 TCGCTGGAAGAGAGAGAGCTGATCCGCGAAGGCGCGCGCTGCTGTAAGCTTGGCGCGCG 540
DB 481 TCGCTGGAAGAGAGAGAGCTGATCCGCGAAGGCGCGCGCTGCTGTAAGCTTGGCGCGCG 540
QY 541 GCGTCGCGCGCTGCTGATCAAGCGCGCGCTGCGCGAGCTGATGAGTGGCGCAACCGC 600
DB 541 GCGTCGCGCGCTGCTGATCAAGCGCGCGCTGCGCGAGCTGATGAGTGGCGCAACCGC 600
QY 601 AATGCGATGATCCGCGAGATCGCAATGCTTCCCTGCTGGAGCTGATGAGTGGCACTGCG 660
DB 601 AATGCGATGATCCGCGAGATCGCAATGCTTCCCTGCTGGAGCTGATGAGTGGCACTGCG 660
QY 661 ACCTGCTTCAAGTCGCGGATCAACACGAGCGCGCGCAATCGCTGCAACGCGCGATC 720
DB 661 ACCTGCTTCAAGTCGCGGATCAACACGAGCGCGCGCAATCGCTGCAACGCGCGATC 720
QY 721 GTGCAATCCGCGCGCATCTCTTGGCTGCAACCTTCCGATGATCTTGGCTGCTGCAAC 780
DB 721 GTGCAATCCGCGCGCATCTCTTGGCTGCAACCTTCCGATGATCTTGGCTGCTGCAAC 780
QY 781 GCGCTGAGAGCGAGCTGCTTCTGCGACATGTGATGATGAGTGGCGCGCTGCACTTGG 840
DB 781 GCGCTGAGAGCGAGCTGCTTCTGCGACATGTGATGATGAGTGGCGCGCTGCACTTGG 840
QY 841 AAGAACGTGCGCGCTGATCGCGCGCGCTGAGCTGATCAAGCGCGCGCGCTGCAAG 900
DB 841 AAGAACGTGCGCGCTGATCGCGCGCGCTGAGCTGATCAAGCGCGCGCGCTGCAAG 900
QY 901 GACATGCGCATCGAGCTCAACGAGATGTACCGGAGTGGGAGCTTGGTGAAGTACCGCTCC 960
DB 901 GACATGCGCATCGAGCTCAACGAGATGTACCGGAGTGGGAGCTTGGTGAAGTACCGCTCC 960
QY 961 TTGCGGTATGAGCGCACTCTTGGCGGTGCTGCACTTACGAGTGGCGAGCGCGCGGCTG 1020
DB 961 TTGCGGTATGAGCGCACTCTTGGCGGTGCTGCACTTACGAGTGGCGAGCGCGCGGCTG 1020
QY 1021 GAGCTGCGGAGGATCGACACGAGCTGAGCGCGCGCGCATGTGCTTCCATGAGCGCG 1080
DB 1021 GAGCTGCGGAGGATCGACACGAGCTGAGCGCGCGCGCATGTGCTTCCATGAGCGCG 1080
QY 1081 ATGCTGATGCTGCGCGAGGAGGAGTGGCGGTGCGCGCGGATGATGCGAGCGACATCTTG 1140
DB 1081 ATGCTGATGCTGCGCGAGGAGGAGTGGCGGTGCGCGCGGATGATGCGAGCGACATCTTG 1140
QY 1141 ATGCTGAGGAGGAGGAGTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200
DB 1141 ATGCTGAGGAGGAGGAGTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200
QY 1201 ATCATTCGCGAAC 1212
DB 1201 ATCATTCGCGAAC 1212

```

```

RESULT 8
ACC69514
ID ACC69514 standard; DNA; 1215 BP.
XX
AC ACC69514;
XX
DT 21-JUL-2003 (first entry)
XX
DE Erwinia sp. (DSM 97-934) creatinase encoding DNA SEQ ID NO:1.
XX
KM Erwinia; creatinase; creatine amidinohydrolase; enzyme; EC 3.5.3.3;
XX Erwinia-type creatinase; creatine; creatinine; uraemia; gigantism;
XX chronic nephritis; acute nephritis; tonic muscular dystrophy; gene; ds.
XX
OS Erwinia sp.
XX
FH Key Location/Qualifiers
FT CDS 1..1215
FT /tag= a
FT /EC_number= "3.5.3.3"
FT /Product= "creatinase"
XX
PD EP1298213-A1.
XX
PF 17-SEP-2002; 2002EP-00020793.
XX
PR 20-SEP-2001; 2001EP-00121780.
XX
PA (HOPE) ROCHE DIAGNOSTICS GMBH.
PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
XX
PI Shao Z, Schmuck R, Kratzsch P, Kenklies J, Weisser H;
PI WPI; 2003-383634/37.
XX
DR P-PSDB; ABR43467.
XX
DR
XX
PT New variant of an Erwinia-type creatinase modified relative to a wild-
PT type creatinase having creatinase activity, useful for determining
PT creatinase and/or creatine concentration in a sample.
XX
PS Disclosure; Page 15-17; 51pp; English.
XX
CC The present invention describes a variant of an Erwinia-type creatinase
CC (I) modified relative to a wild-type creatinase having creatinase
CC activity. The variant comprises at least one amino acid substitution at a
CC position of the fully defined 404 amino acid sequence given in ABR43467;
CC these are selected from N130, M203, I278, I304 and F395. Creatinase has
CC the EC number EC 3.5.3.3 and is also known as creatine amidinohydrolase.
CC Also described is a reagent (II) for determining creatine comprising the
CC Erwinia-type creatinase variant. The variant is useful for determining
CC creatinine and/or creatine concentration in a sample. Measuring
CC creatinine and/or creatine are useful for diagnosing uraemia, chronic
CC nephritis, acute nephritis, gigantism, tonic muscular dystrophy and some
CC other related diseases. The mutant enzymes have improved stability, lower
CC conductivity and/or lower Km-values for creatine; they are much better
CC suited to detection methods for creatine. The present sequence encodes
CC Erwinia sp. (DSM 97-934) wild-type creatinase from the present invention
XX
SO Sequence 1215 BP; 242 A; 392 C; 358 G; 223 T; 0 U; 0 Other;
Query Match 81.3%; Score 984.8; DB 8; Length 1215;
Best Local Similarity 88.3%; Pred. No. 6,8e-158;
Matches 1070; Conservative 0; Mismatches 142; Indels 0; Gaps 0;
QY 1 AAGACTGAGAGATGTCAGTGTGATGAAATGAGCAAGGCGGAGAAATATTTCGCGC 60
DB 1 AAGACTGAGAGATGTCAGTGTGATGAAATGAGCAAGGCGGAGAAATATTTCGCGC 60
QY 61 TTTTCGATGCGGAGATGACCGCGCGCAAAAGAGCTTGGCGGCTGATGCGCAAGAAC 120
DB 61 TTTTCGATGCGGAGATGACCGCGCGCAAAAGAGCTTGGCGGCTGATGCGCAAGAAC 120

```

```

Db      61  TTTCCGATGCCAGATGACGCCGCCGACAGTGAAGTGGCGGCTGGATGGCCGAAC 120
Qy      121  AATGTCATGCGCGCGCTGTTCACCTCTTATCATGTCATCACTACTATTCGGCTGGCTG 180
Db      121  GAGCTGAGCGCTGCGCTGTTCACCTCTTATCATGTCATCACTACTATTCGGCTGGCTG 180
Qy      181  TACTGCTATTTGGCGCGCAATAGCGATGTCATGACGACGACCATGCGACGACATC 240
Db      181  TACTGCTATTTGGCGCGCAATAGCGATGTCATGACGACGACCATGCGACGACATC 240
Qy      241  TCGCGCGCATGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
Db      241  TCGCGCGCATGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
Qy      301  ACCGACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
Db      301  ACCGACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
Qy      361  CGCATCGCATGACGATTCGACACGCTCAATCTGACCTTCGCGCGCGCGCGCGCGCGCG 420
Db      361  CGCATCGCATGACGATTCGACACGCTCAATCTGACCTTCGCGCGCGCGCGCGCGCGCG 420
Qy      421  CTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
Db      421  CTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
Qy      481  TCGCTCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Db      481  TCGCTCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Qy      541  GCGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
Db      541  GCGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
Qy      601  AATGCGATGATCCGCGAGATTCGCAATGCTTCCTTCGCGCGCGCGCGCGCGCGCGCG 660
Db      601  AATGCGATGATCCGCGAGATTCGCAATGCTTCCTTCGCGCGCGCGCGCGCGCGCGCG 660
Qy      661  ACTGCTTCGATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
Db      661  ACTGCTTCGATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
Qy      721  GTGCAATCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 780
Db      721  GTGCAATCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 780
Qy      781  GCGCTGAGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 840
Db      781  GCGCTGAGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 840
Qy      841  AAGAACGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
Db      841  AAGAACGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
Qy      901  GACATGCGCATGAGCTCAACGAGATGTCACGCGAGTGGGAACTGCTGAAATCCGCTCC 960
Db      901  GACATGCGCATGAGCTCAACGAGATGTCACGCGAGTGGGAACTGCTGAAATCCGCTCC 960
Qy      961  TTGCGCTATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1020
Db      961  TTGCGCTATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1020
Qy      1021  GAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1080
Db      1021  GAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1080
Qy      1081  AATGCGATGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1140
Db      1081  AATGCGATGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1140
Qy      1141  ATGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1200
Db      1141  ATGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1200

```

```

Qy      1201  ATCATCCGCAAC 1212
Db      1201  ATCATCCGCAAC 1212

RESULT 9
ACC69519
ID  ACC69519 standard; DNA; 1215 BP.
XX
AC  ACC69519;
XX
DT  21-JUL-2003 (first entry)
XX
DE  Mutant Erwinia creatinase CTgc2 encoding DNA SEQ ID NO:17.
XX
KW  Erwinia; creatinase; creatine amidinohydrolase; enzyme; EC 3.5.3.3;
KW  Erwinia-type creatinase; creatine; creatinine; uraemia; gigantism;
KW  chronic nephritis; acute nephritis; tonic muscular dystrophy; mutant;
KW  gene; ds.
XX
OS  Erwinia sp.
OS  Synthetic.
XX
FH  Key
FT  CDS
FT  1.1215
FT  /tag= a
FT  /EC_number= "3.5.3.3"
FT  /product= "creatinase mutant CTgc2"
XX
PN  EP1298213-A1.
XX
PD  02-APR-2003.
XX
PF  17-SEP-2002; 2002EP-00020793.
XX
PR  20-SEP-2001; 2001EP-00121780.
XX
PA  (HOPE ) ROCHE DIAGNOSTICS GMBH.
PA  (HOPE ) HOFFMANN LA ROCHE & CO AG F.
XX
PI  Shao Z, Schmuck R, Kratzsch P, Kenkies J, Weisser H;
XX
DR  MPI; 2003-383834/37.
XX
DR  P-PDB; ABR43478.
XX
PT  New variant of an Erwinia-type creatinase modified relative to a wild-
PT  type creatinase having creatinase activity, useful for determining
PT  creatinine and/or creatine concentration in a sample.
XX
PS  Example 4; Page 33-35; 51pp; English.
XX
CC  The present invention describes a variant of an Erwinia-type creatinase
CC  (I) modified relative to a wild-type creatinase having creatinase
CC  activity. The variant comprises at least one amino acid substitution at a
CC  position of the fully defined 404 amino acid sequence given in ABR43467;
CC  these are selected from N130, M203, I278, I1304 and P395. Creatinase has
CC  the EC number EC 3.5.3.3 and is also known as creatine amidinohydrolase.
CC  Also described is a reagent (II) for determining creatine comprising the
CC  Erwinia-type creatinase variant. The variant is useful for determining
CC  creatinine and/or creatine concentration in a sample. Measuring
CC  creatinine and creatine are useful for diagnosing uraemia, chronic
CC  nephritis, acute nephritis, gigantism, tonic muscular dystrophy and some
CC  other related diseases. The mutant enzymes have improved stability, lower
CC  conductivity and/or lower Km-values for creatine; they are much better
CC  suited to detection methods for creatine. The present sequence encodes a
CC  mutant Erwinia creatinase from the present invention
XX
SQ  Sequence 1215 BP; 241 A; 392 C; 361 G; 221 T; 0 U; 0 Other;

```

```

Query Match      81.3%; Score 984.8; DB 8; Length 1215;
Best Local Similarity 88.3%; Pred. No. 6,8e-158;
Matches 1070; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

```

QY 1 ATGACTGACGACATGTTGACCTGATGAATGACACAGCGGAGAAATATTGCGCG 60
 Db 1 ATGACTGACGACATGTTGACCTGATGAATGACACAGCGGAGAAATATTGCGCG 60
 QY 61 TTTTGGGATGCGGAGATGACCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
 Db 61 TTTTGGGATGCGGAGATGACCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
 QY 121 AATGTCATGCGGCGCTGTTCACTCTTATCACTGTCATCACTACTATTTCGCGT 180
 Db 121 AATGTCATGCGGCGCTGTTCACTCTTATCACTGTCATCACTACTATTTCGCGT 180
 QY 181 TACTGCTATTTGACAGCGCAATGCGCATGCTGATGACCAACAACCGCGACGATT 240
 Db 181 TACTGCTATTTGACAGCGCAATGCGCATGCTGATGACCAACAACCGCGACGATT 240
 QY 241 TGGGCGGCGATGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 300
 Db 241 TGGGCGGCGATGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 300
 QY 301 ACCGACTGCGCGCGCGACAAATTTCTATCGCGCGTGCACAGCTGACACAGCGCG 360
 Db 301 ACCGACTGCGCGCGCGACAAATTTCTATCGCGCGTGCACAGCTGACACAGCGCG 360
 QY 361 GCGATGCGCATGAGTTGACACAGCTGATCTGACCTTCCGCGCGGCGGCGGCGG 420
 Db 361 GCGATGCGCATGAGTTGACACAGCTGATCTGACCTTCCGCGCGGCGGCGGCGG 420
 QY 421 CTACCGGCGGCTGAGTTGCTTCAACATGCGCGGCGGCGGCGGCGGCGGCGGCG 480
 Db 421 CTACCGGCGGCTGAGTTGCTTCAACATGCGCGGCGGCGGCGGCGGCGGCGGCG 480
 QY 481 TCGCTGGAAGACAGAGCTGATCCGCGAGGCGCGCGGCTGATGAGCTGCGCGCG 540
 Db 481 TCGCTGGAAGACAGAGCTGATCCGCGAGGCGCGCGGCTGATGAGCTGCGCGCG 540
 QY 541 GCGTGGCGGCGGCTGATCAAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 600
 Db 541 GCGTGGCGGCGGCTGATCAAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 600
 QY 601 AATGCGATGATCGCGAGATCGCGAAATGCTTCCCTTGGGAGCTGATGAGCACT 660
 Db 601 AATGCGATGATCGCGAGATCGCGAAATGCTTCCCTTGGGAGCTGATGAGCACT 660
 QY 661 ACCGCTTTCAGTCCGCGCATCAACACGAGCGCGCGGCGGCGGCGGCGGCGG 720
 Db 661 ACCGCTTTCAGTCCGCGCATCAACACGAGCGCGCGGCGGCGGCGGCGGCGG 720
 QY 721 GTGCAATCGCGGAGCATCTTTCGCTCAACACTTCCGATGATCTTTCGCTTAC 780
 Db 721 GTGCAATCGCGGAGCATCTTTCGCTCAACACTTCCGATGATCTTTCGCTTAC 780
 QY 781 GGGCTGAGAGCGACGCTTTCGCGCAATGCTGATGAGCGGCGGCGGCGGCGG 840
 Db 781 GGGCTGAGAGCGACGCTTTCGCGCAATGCTGATGAGCGGCGGCGGCGGCGG 840
 QY 841 AAGAACTGCGCGGATGAGCGCGCGGCTGAGCTGATCAAGCGCGCGCGGCTGCA 900
 Db 841 AAGAACTGCGCGGATGAGCGCGCGGCTGAGCTGATCAAGCGCGCGCGGCTGCA 900
 QY 901 GACATGCGCATGAGCTCAACAGAGATGATCCGCGAGTGGAGCTGCTGAAATG 960
 Db 901 GACATGCGCATGAGCTCAACAGAGATGATCCGCGAGTGGAGCTGCTGAAATG 960
 QY 961 TTCGCTATGCGGCACTCTTTCGCGGCTGCTGCTTCCACTACTACGCGCGGCG 1020
 Db 961 TTCGCTATGCGGCACTCTTTCGCGGCTGCTGCTTCCACTACTACGCGCGGCG 1020
 QY 1021 GAGCTGCGGAGGACATGACACCGAGCTGAGACCGCGGCGGCGGCGGCGGCGG 1080
 Db 1021 GAGCTGCGGAGGACATGACACCGAGCTGAGACCGCGGCGGCGGCGGCGGCGG 1080

QY 1081 ATGCGATGCTGCGGAGGAGCATGCGCGGCTGCGGCGGCTGATGCGGAGACATCTCG 1140
 Db 1081 ATGCGATGCTGCGGAGGAGCATGCGCGGCTGCGGCGGCTGATGCGGAGACATCTCG 1140
 QY 1141 ATCGTGGGAGGAGCGGTGCGGAGACATCAACCGGCTTCCGTTGCGGTCCGGAACAAC 1200
 Db 1141 ATCGTGGGAGGAGCGGTGCGGAGACATCAACCGGCTTCCGTTGCGGTCCGGAACAAC 1200
 QY 1201 ATCATCCGCAAC 1212
 Db 1201 ATCATCCGCAAC 1212
 RESULT 10
 ACC69515
 ID ACC69515 standard; DNA; 1215 BP.
 AC C69515;
 DT 21-JUL-2003 (first entry)
 DE Mutant Erwinia creatinase CTIm24 encoding DNA SEQ ID NO:9.
 KW Erwinia; creatinase; creatine amidinohydrolyase; enzyme; EC 3.5.3.3;
 KW Erwinia-type creatinase; creatine; creatinine; uremia; gigantism;
 KW chronic nephritis; acute nephritis; tonic muscular dystrophy; mutant;
 KW gene; ds.
 OS Erwinia sp.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT 1..1215
 FT /*tag= a
 FT /EC number= "3.5.3.3"
 FT /product= "creatinase mutant CTIm24"
 PN EP1298213-A1.
 PD 02-APR-2003.
 PF 17-SEP-2002; 2002EP-00020793.
 PR 20-SEP-2001; 2001EP-00121780.
 PA (HOFF) ROCHE DIAGNOSTICS GMBH.
 PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
 PI Shao Z, Schmuck R, Kratzsch P, Kenkies J, Weisner H;
 DR WPI; 2003-383834/37.
 DR P-PSDB; ABR43474.
 XX New variant of an Erwinia-type creatinase modified relative to a wild-
 PT type creatinase having creatinase activity, useful for determining
 PT creatinine and/or creatine concentration in a sample.
 XX
 PS Example 4; Page 20-21; 51pp; English.
 CC The present invention describes a variant of an Erwinia-type creatinase
 CC (I) modified relative to a wild-type creatinase having creatinase
 CC activity. The variant comprises at least one amino acid substitution at a
 CC position of the fully defined 404 amino acid sequence given in ABR43467;
 CC these are selected from N130, M203, I278, I1304 and F355. Creatinase has
 CC the EC number EC 3.5.3.3 and is also known as creatine amidinohydrolyase.
 CC Also described is a reagent (II) for determining creatine comprising the
 CC Erwinia-type creatinase variant. The variant is useful for determining
 CC creatinine and/or creatine concentration in a sample. Measuring
 CC creatinine and creatine are useful for diagnosing uraemia, chronic
 CC nephritis, acute nephritis, gigantism, tonic muscular dystrophy and some
 CC other related diseases. The mutant enzymes have improved stability, lower
 CC conductivity and/or lower Km-values for creatine; they are much better
 CC suited to detection methods for creatine. The present sequence encodes a

the EC number EC 3.5.3.3 and is also known as creatine amidinohydrolase. Also described is a reagent (II) for determining creatine comprising the Erwinia-type creatinase variant. The variant is useful for determining creatinine and/or creatine concentration in a sample. Measuring creatinine and creatine are useful for diagnosing uraemia, chronic nephritis, acute nephritis, gigantism, tonic muscular dystrophy and some other related diseases. The mutant enzymes have improved stability, lower conductivity and/or lower Km-values for creatine; they are much better suited to detection methods for creatine. The present sequence encodes a mutant Erwinia creatinase from the present invention

Sequence 1212 BP; 240 A; 394 C; 358 G; 220 T; 0 U; 0 Other;

Query Match 80.3%; Score 980; DB 8; Length 1212;
Best Local Similarity 88.0%; Pred. No. 4,4e-157;
Matches 1067; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

QY 1 ATGACTGACGACATGTTGACGATGATGAATGGACAAACGGGAGAAAGATTATTCGCGG 60
DB 1 ATGACTGACGACATGTTGACGATGATGAATGGACAAATGTAGAAAGAAATTATTCGCC 60
QY 61 TTTTCGATGCGGACGATGACCCCGCCCAACGACGTTGCGCGCTGATGGCCAAAGAC 120
DB 61 TTTTCGATGCGGACGATGACCGCGCCCAAGTGAAGCGTGGCGCTGATGGCCAAAGAC 120
QY 121 AATGTGATGCGGCGCTGTTCACTCTTATCACTGATCACTAATCTATTCGGCTGCTG 180
DB 121 GACGTGACGCTGCGCTGTTCACTCTTATCACTGATCACTAATCTATTCGGCTGCTG 180
QY 181 TACTGCTATTTGGAGACGACAGTACGGGATGATGACACCAACGCGACAGCAATT 240
DB 181 TACTGCTATTTGGAGACGACAGTACGGGATGATGACACCAACGCGACAGCAATT 240
QY 241 TCGGCGGACATCGACCGCGGACAGCCCTGCGCGGACGCTTGGGACAAATCACTAC 300
DB 241 TCGGCGGACATCGACCGCGGACAGCCCTGCGCGGACGCTTGGGACAAATCACTAC 300
QY 301 ACCGATCGGCGCGGACCAATTTCTATCGCGCGCTGCGGACGCTGACCAACGCGCG 360
DB 301 ACCGATCGGCGCGGACCAATTTCTATCGCGCGCTGCGGACGCTGACCAACGCGCG 360
QY 361 CGCATCGGACATCGAGTTCGACACGTCATCTGCACTTCGCGCGGACGCTGACGAAG 420
DB 361 CGCATCGGACATCGAGTTCGACACGTCATCTGCACTTCGCGCGGACGCTGACGAAG 420
QY 421 CTACCGGCGCTCGAGTTCGTCGACATCGACGCGCTCGATGTGATGCGCACATCAAG 480
DB 421 CTACCGGCGCTCGAGTTCGTCGACATCGACGCGCTCGATGTGATGCGCACATCAAG 480
QY 481 TCGCTGGAAGACGAAGCTGATTCGCGAAGCGCGCGCTGTGTGACGTCGCGCGCGG 540
DB 481 TCGCTGGAAGACGAAGCTGATTCGCGAAGCGCGCGCTGTGTGACGTCGCGCGCGG 540
QY 541 GCGTGGCGGCTGCGATCAAGCGCGGCGTCCGCGACGATGAAGTGGGATCGCACCA 600
DB 541 GCGTGGCGGCTGCGATCAAGCGCGGCGTCCGCGACGATGAAGTGGGATCGCACCA 600
QY 601 AATGCAATGATCGCGGACATCGCCAAATCGTTCCCTTGTGAGCTGATGACACCTGG 660
DB 601 AATGCAATGATCGCGGACATCGCCAAATCGTTCCCTTGTGAGCTGATGACACCTGG 660
QY 661 ACCGTGTTCCAGTGGGACATCAACCGAGCGCGGACAAATCGGAGAACCGGATC 720
DB 661 ACCGTGTTCCAGTGGGACATCAACCGAGCGCGGACAAATCGGAGAACCGGATC 720
QY 721 GTGCAATCGGCGACATCTTGTGCTCAACACCTTCCGATGATCTTCCGCTACTAC 780
DB 721 GTGCAATCGGCGACATCTTGTGCTCAACACCTTCCGATGATCTTCCGCTACTAC 780
QY 781 GCGCTGAGGACGACGCTGTTCTGCGACCATGTGATGACCGACCTGACATCTGGAG 840
DB 781 GCGCTGAGGACGACGCTGTTCTGCGACCATGTGATGACCGACCTGACATCTGGAG 840

QY 841 AAGACGTGCGCGTGCATCGCGCGGCTCGAGCTGATCAAGCGGCGCGCTGCAAG 900
DB 841 AAGACGTGCGCGTGCATCGCGCGGCTCGAGCTGATCAAGCGGCGCGCTGCAAG 900
QY 901 GACATGCGCATGAGCTCAACGATGTAACCGAGTGGACCTGTAAGTAACCGCTCC 960
DB 901 GATATGCGCATGAGCTCAACGATGTAACCGAGTGGAGATGCTGTAAGTAACCGCTCC 960
QY 961 TTGCGTATGCGGACCTCTTCCGCGCTGCTGCTGCTACTACTACGCTGCGAGCGGCG 1020
DB 961 TTGCGTATGCGGACCTCTTCCGCGCTGCTGCTGCTACTACTACGCTGCGAGCGGCG 1020
QY 1021 GAGCTGCGGAGACATCGACCGAGCTGAAGCCGCGGATGCTGCTCAATGAGCGG 1080
DB 1021 GAGCTGCGGAGACATCGATTCATTCGCTGCAACCGCGGATGCTGCTCAATGAGCGG 1080
QY 1081 ATGCTGATGCTGCGGAGGACATGCGGCTGCGGCTGCTATGCGACGACATCTG 1140
DB 1081 ATGCTGATGCTGCGGAGGACATGCGGCTGCGGCTGCTATGCGACGACATCTG 1140
QY 1141 ATGCTGCGGAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
DB 1141 ATGCTGCGGAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200
QY 1201 ATCATCGGCAAC 1212
DB 1201 ATCATCGGCAAC 1212

RESULT 12
ID ACC69516 standard; DNA; 1212 BP.
XX AC ACC69516;
XX DT 21-JUL-2003 (first entry)
XX DE Mutant Erwinia creatinase CT2m9 encoding DNA SEQ ID NO.11.
XX KW Erwinia; creatinase; creatine amidinohydrolase; enzyme; EC 3.5.3.3;
XX KM Erwinia-type creatinase; creatine; creatinine; uraemia; gigantism;
XX KM chronic nephritis; acute nephritis; tonic muscular dystrophy; mutant;
XX KM gene; de.
OS Erwinia sp.
OS Synthetic.
OS
FH Key Location/Qualifiers
FT CDS 1..1212
FT FT /tag= a
FT FT /partial
FT FT /EC_number= "3.5.3.3"
FT FT /product= "creatinase mutant CT2m9"
FT FT /note= "no stop codon given"
XX
XX PN EP1298213-A1.
XX PD 02-APR-2003.
XX PF 17-SEP-2002; 2002BP-00020793.
XX PR 20-SEP-2001; 2001BP-00121780.
XX PA (HOFF) ROCHE DIAGNOSTICS GMBH.
XX PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
XX PI Shao Z, Schmuck R, Kratzsch P, Kenklee J, Weisser H;
XX DR WPI; 2003-383834/37.
XX DR P-PDB; ABR3475.
XX PT New variant of an Erwinia-type creatinase modified relative to a wild-type creatinase having creatinase activity, useful for determining

PT creatinine and/or creatine concentration in a sample.
 XX
 PS Example 4; Page 23-25; 51pp; English.
 XX
 CC The present invention describes a variant of an Erwinia-type creatinase
 CC (I) modified relative to a wild-type creatinase having creatinase
 CC activity. The variant comprises at least one amino acid substitution at a
 CC position of the fully defined 404 amino acid sequence given in ABR43467:
 CC these are selected from N130, M203, I278, I1304 and P395. Creatinase has
 CC the EC number EC.3.5.3.3 and is also known as creatine amidinohydrolase.
 CC Also described is a reagent (II) for determining creatine comprising the
 CC Erwinia-type creatinase variant. The variant is useful for determining
 CC creatinine and/or creatine concentration in a sample. Measuring
 CC creatinine and creatine are useful for diagnosing uraemia, chronic
 CC nephritis, acute nephritis, gigantism, tonic muscular dystrophy and some
 CC other related diseases. The mutant enzymes have improved stability, lower
 CC conductivity and/or lower Km-values for creatine; they are much better
 CC suited to detection methods for creatine. The present sequence encodes a
 CC mutant Erwinia creatinase from the present invention
 XX
 SQ Sequence 1212 BP; 239 A; 393 C; 359 G; 221 T; 0 U; 0 Other;
 Query Match 80.9%; Score 980; DB 8; Length 1212;
 Best Local Similarity 88.0%; Pred. No. 4,4e-157;
 Matches 1067; Conservative 0; Mismatches 145; Indels 0; Gaps 0;
 QY 1 ATGACGTGACGACATGTTGACGTGATGAAATGGACACACGCGAGAAAGATTATTCGCGC 60
 DB 1 ATGACGTGACGACATGTTGACGTGATGAAATGGACACACGCGAGAAATATTCGCGC 60
 QY 61 TTTTCGAGATGCCGAGATGACCCGCGCCCAAGACAGCTTCGCGATGAGCCCAAGAC 120
 DB 61 TTTTCGAGATGCCGAGATGACCCGCGCCCAAGACAGCTTCGCGATGAGCCCAAGAC 120
 QY 121 AATGTCGATGCGCGCGCTGTTCACTTATCACTGATCACTACTATTCGCGCTG 180
 DB 121 AATGTCGATGCGCGCGCTGTTCACTTATCACTGATCACTACTATTCGCGCTG 180
 QY 121 GACGTGACGCTGCGCTGTTCACTTATCACTGATCACTACTATTCGCGCTG 180
 DB 121 GACGTGACGCTGCGCTGTTCACTTATCACTGATCACTACTATTCGCGCTG 180
 QY 181 TACTGCTATTTGGAGAGCAAGTAACGGATGATGATGACACAAACGCGACAGAT 240
 DB 181 TACTGCTATTTGGAGAGCAAGTAACGGATGATGATGACACAAACGCGACAGAT 240
 QY 181 TACTGCTATTTGGAGAGCAAGTAACGGATGATGATGACACAAACGCGACAGAT 240
 DB 181 TACTGCTATTTGGAGAGCAAGTAACGGATGATGATGACACAAACGCGACAGAT 240
 QY 241 TCGGCGCGGATCGACCGCGCGCGAGCCCTGCGCGCGAGCTTCGCGCAACATCACTAC 300
 DB 241 TCGGCGCGGATCGACCGCGCGCGAGCCCTGCGCGCGAGCTTCGCGCAACATCACTAC 300
 QY 241 TCGGCGCGGATCGACCGCGCGCGAGCCCTGCGCGCGAGCTTCGCGCAACATCACTAC 300
 DB 241 TCGGCGCGGATCGACCGCGCGCGAGCCCTGCGCGCGAGCTTCGCGCAACATCACTAC 300
 QY 301 ACCGACTGCGCGCGCGCAATTTCTATCGCGCGCGTGGCGAGCTGACCAACGCGCGCAAG 360
 DB 301 ACCGACTGCGCGCGCGCAATTTCTATCGCGCGCGTGGCGAGCTGACCAACGCGCGCAAG 360
 QY 301 ACCGACTGCGCGCGCGCAATTTCTATCGCGCGCGTGGCGAGCTGACCAACGCGCGCAAG 360
 DB 301 ACCGACTGCGCGCGCGCAATTTCTATCGCGCGCGTGGCGAGCTGACCAACGCGCGCAAG 360
 QY 361 CGGATGCGATCGATGATTCGACACGTCATCTGACTTCGCGCGCGAGCTCGAAGAGCGC 420
 DB 361 CGGATGCGATCGATGATTCGACACGTCATCTGACTTCGCGCGCGAGCTCGAAGAGCGC 420
 QY 361 CGGATGCGATCGATGATTCGACACGTCATCTGACTTCGCGCGCGAGCTCGAAGAGCGC 420
 DB 361 CGGATGCGATCGATGATTCGACACGTCATCTGACTTCGCGCGCGAGCTCGAAGAGCGC 420
 QY 421 CTACCGGCGGCTGAGTTCTGCGACATCAAGCCCTCGATGATGATGAGCGCACATCAAG 480
 DB 421 CTACCGGCGGCTGAGTTCTGCGACATCAAGCCCTCGATGATGATGAGCGCACATCAAG 480
 QY 421 CTACCGGCGGCTGAGTTCTGCGACATCAAGCCCTCGATGATGATGAGCGCACATCAAG 480
 DB 421 CTACCGGCGGCTGAGTTCTGCGACATCAAGCCCTCGATGATGATGAGCGCACATCAAG 480
 QY 481 TCGCTCGAAGAGCAAGTGTATCGCGAGAGCGCGCGCTGTGTGACGTTCGCGCGCGC 540
 DB 481 TCGCTCGAAGAGCAAGTGTATCGCGAGAGCGCGCGCTGTGTGACGTTCGCGCGCGC 540
 QY 481 TCGCTCGAAGAGCAAGTGTATCGCGAGAGCGCGCGCTGTGTGACGTTCGCGCGCGC 540
 DB 481 TCGCTCGAAGAGCAAGTGTATCGCGAGAGCGCGCGCTGTGTGACGTTCGCGCGCGC 540
 QY 541 GCGTGGCGGCTGCGATCAAGCGCGCGTGCAGAGCATGAGTGGCATCGCGACAC 600
 DB 541 GCGTGGCGGCTGCGATCAAGCGCGCGTGCAGAGCATGAGTGGCATCGCGACAC 600
 QY 541 GCGTGGCGGCTGCGATCAAGCGCGCGTGCAGAGCATGAGTGGCATCGCGACAC 600
 DB 541 GCGTGGCGGCTGCGATCAAGCGCGCGTGCAGAGCATGAGTGGCATCGCGACAC 600
 QY 601 AATGCGATGATCGCGAGATCGCAATGTTCCCTTCGCGAGCGATGAGCACTCG 660
 DB 601 AATGCGATGATCGCGAGATCGCAATGTTCCCTTCGCGAGCGATGAGCACTCG 660
 QY 601 AATGCGATGATCGCGAGATCGCAATGTTCCCTTCGCGAGCGATGAGCACTCG 660
 DB 601 AATGCGATGATCGCGAGATCGCAATGTTCCCTTCGCGAGCGATGAGCACTCG 660
 QY 661 ACCGCTTCGATCGCGGATCAACACCGAGCGCGCGCATTCGCTGACCAACCGCATC 720
 DB 661 ACCGCTTCGATCGCGGATCAACACCGAGCGCGCGCATTCGCTGACCAACCGCATC 720

QY 721 GTGCAATCGCGAGACATTCCTTTGCTCAACACCTTCCGATGATCTTCGCTACTACACC 780
 DB 721 GTGCAATCGCGAGATTCCTTCGTCTCAACACCTTCCGATGATCTTCGCTACTACACC 780
 QY 781 GCGCTGAGAGCGACGCTGTTCTGCGACATGTCAGTACGCGACCTCGCATCTGGAG 840
 DB 781 GCGCTGAGAGCGAAGCGCTGTTCTGCGACATGTCAGTACGCGACCTCGCATCTGGAG 840
 QY 841 AAGAAAGTGGCGGTGATGATGCGCGCGGCTGAGCTGATCAAGCGCGCGCGCTGCAAG 900
 DB 841 AAGAAAGTGGCGGTGATGATGCGCGCGGCTGAGCTGATCAAGCGCGCGCGCTGCAAG 900
 QY 901 GACATGCGCATCGAGCTCAACAGATGTAACCGAGTGGAGCCTGCTGAACTACCGCTCC 960
 DB 901 GATATGCGCATCGAATCTCAACAGATGTAACCGAGTGGAGTGTGCTGAACTACCGCTCC 960
 QY 961 TTCGCTATGCGCATCTCTTCGCGGTGCTGTCGCACTACTACGATGCGAGCGCGCTG 1020
 DB 961 TTCGCTATGCGCATCTCTTCGCGGTGCTGTCGCACTACTACGATGCGAGCGCGCTG 1020
 QY 1021 GAGCTGCGGAGAGCATGACACCGAGCTGAAGCGCGGATGATGATGATGAGCGCG 1080
 DB 1021 GAGCTGCGGAGAGCATGACACCGAGCTGAAGCGCGGATGATGATGATGAGCGCG 1080
 QY 1081 ATGATGATGCTGCGCGAGAGCATGCGCGGTGCGCGCGGCTATGCGAGACGACATCTCT 1140
 DB 1081 ATGATGATGCTGCGCGAGAGCATGCGCGGTGCGCGCGGCTATGCGAGACGACATCTCT 1140
 QY 1141 ATGCTGCGGAGAGAGCGGTGCGCGAGACATCAACCGCTTCGCTGCGAGACCAAC 1200
 DB 1141 ATGCTGCGGAGAGAGCGGTGCGCGAGACATCAACCGGATTCGCTGCGAGACCAAC 1200
 QY 1201 ATCATCCGCAAC 1212
 DB 1201 ATCATCCGCAAC 1212

RESULT 13
 ACCG9518
 ID ACCG9518 standard; DNA; 1215 BP.
 XX
 AC ACCG9518;
 XX
 DT 21-JUL-2003 (first entry)
 XX
 DE Mutant Erwinia creatinase CT2m28 encoding DNA SEQ ID NO:15.
 XX
 KW Erwinia; creatinase; creatine amidinohydrolase; enzyme; EC 3.5.3.3;
 KW Erwinia-type creatinase; creatine; creatinine; uraemia; gigantism;
 KW chronic nephritis; acute nephritis; tonic muscular dystrophy; mutant;
 KW gene; de.
 XX
 OS Erwinia sp.
 OS Synthetic.
 XX
 PH Key
 FT CDS
 FT Location/Qualifiers
 FT 1..1215
 FT /tag= a
 FT /EC number= "3.5.3.3"
 FT /product= "creatinase mutant CT2m28"
 XX
 PN BP1298213-A1.
 XX
 PD 02-APR-2003.
 XX
 PF 17-SEP-2002; 2002BP-00020793.
 XX
 PR 20-SEP-2001; 2001BP-00121780.
 XX
 XX (HOFF) ROCHE DIAGNOSTICS GMBH.
 XX (HOFF) HOFFMANN LA ROCHE & CO AG F.
 XX

PI Shao Z, Schmuck R, Kratzsch P, Kenkies J, Weisser H;
 XX WPI: 2003-383834/37.
 DR P-PSDB; ABR34377.
 XX
 PT New variant of an Erwinia-type creatinase modified relative to a wild-
 PT type creatinase having creatinase activity, useful for determining
 PT creatinase and/or creatine concentration in a sample.
 XX
 PS Example 4; Page 30-32; Sipp; English.
 XX
 CC The present invention describes a variant of an Erwinia-type creatinase
 CC (I) modified relative to a wild-type creatinase having creatinase
 CC activity. The variant comprises at least one amino acid substitution at a
 CC position of the fully defined 404 amino acid sequence given in ABR34367.
 CC These are selected from N130, M203, I278, I1304 and F335. Creatinase has
 CC the EC number EC 3.5.3.3 and is also known as creatine amidinohydrolase.
 CC Also described is a reagent (II) for determining creatine comprising the
 CC Erwinia-type creatinase variant. The variant is useful for determining
 CC creatinine and/or creatine concentration in a sample. Measuring
 CC creatinine and creatine are useful for diagnosing uraemia, chronic
 CC nephritis, acute nephritis, gigantism, tonic muscular dystrophy and some
 CC other related diseases. The mutant enzymes have improved stability, lower
 CC conductivity and/or lower Km-values for creatine; they are much better
 CC suited to detection methods for creatine. The present sequence encodes a
 CC mutant Erwinia creatinase from the present invention
 CC
 SQ Sequence 1215 BP; 240 A; 394 C; 359 G; 222 T; 0 U; 0 Other;
 Query Match 80.9%; Score 980; DB 8; Length 1215;
 Best Local Similarity 88.0%; Pred. No. 4,4e-157;
 Matches 1067; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

QY 601 AATGCGATGATCCGCGAGATGCCCAATGCTTCCCTTCGTGAGCTGATGACACCTCG 650
 DB 601 AATGCGATGATGTCGCGAGATGCCCAATGCTTCCCTTCGTGAGCTGATGATGACACCTCG 650
 QY 661 ACCGTGTTCCAGTCGGGCAATCAACCGAGCGCGCAATTCGGTCAACCAACCGCATC 720
 DB 661 ACCGTGTTCCAGTCGGGCAATCAACCGAGCGCGCAATTCGGTCAACCAACCGCATC 720
 QY 721 GTGCAATCGGCGAGATCCCTTCGTGACACCTTCGGATATCTTCGGCTACTACAC 780
 DB 721 GTGCAATCGGCGAGATCCCTTCGTGACACCTTCGGATATCTTCGGCTACTACAC 780
 QY 781 GCGCTGAGCGAGCGCTGTTCTGCGACCAATGTCAGCGGACCTTCACATCTGCGAG 840
 DB 781 GCGCTGAGCGAGCGCTGTTCTGCGACCAATGTCAGCGGACCTTCACATCTGCGAG 840
 QY 841 AAGAACGTGCGCTGATGCTGCGCGGCTGAGCTGATCAAGCGCGCGCGCTGCAAG 900
 DB 841 AAGAACGTGCGCTGATGCTGCGCGGCTGAGCTGATCAAGCGCGCGCGCTGCAAG 900
 QY 901 GATATCGCATGAGGTCAACGAGATGTAACGCGAGTGGGACCTGTAAGTACCGCTCC 960
 DB 901 GATATCGCATGAGGTCAACGAGATGTAACGCGAGTGGGACCTGTAAGTACCGCTCC 960
 QY 961 TTGCGCTATGCGCACTCTTTCGCGGTGCTGTCGCACTACTAGTTCGAGGCGCGCTG 1020
 DB 961 TTGCGCTATGCGCACTCTTTCGCGGTGCTGTCGCACTACTAGTTCGAGGCGCGCTG 1020
 QY 1021 GAGCTGCGCGAGCAATGCAACCGAGCTGAAAGCCCGGATGCTTTCATGAGCGG 1080
 DB 1021 GAGCTGCGCGAGCAATGCAACCGAGCTGAAAGCCCGGATGCTTTCATGAGCGG 1080
 QY 1081 AATGCGATGCTGCGGAGGCGATGCGCGGTGCGCGGCTATCGGAGGACGACATCTG 1140
 DB 1081 AATGCGATGCTGCGGAGGCGATGCGCGGTGCGCGGCTATCGGAGGACGACATCTG 1140
 QY 1141 ATGCTGCGGAGGAGCGGTGCGGAGCAATCAACCGCTTCCGTTGCTCGGAAACAAC 1200
 DB 1141 ATGCTGCGGAGGAGCGGTGCGGAGCAATCAACCGCTTCCGTTGCTCGGAAACAAC 1200
 QY 1201 ATCATCCGCAAC 1212
 DB 1201 ATCATCCGCAAC 1212

RESULT 14
 ACC69521
 ID ACC69521 standard; DNA; 1212 BP.
 XX
 AC ACC69521;
 XX
 DT 21-JUL-2003 (first entry)
 XX
 DE Mutant Erwinia creatinase Ctsd7 encoding DNA SEQ ID NO:21.
 XX
 KW Erwinia; creatinase; creatine amidinohydrolase; enzyme; EC 3.5.3.3;
 KW Erwinia-type creatinase; creatine; creatinine; uraemia; gigantism;
 KW chronic nephritis; acute nephritis; tonic muscular dystrophy; mutant;
 KW gene; ds.
 OS Erwinia sp.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT 1..1212
 FT /*tag= a
 FT /partial
 FT /EC number= "3.5.3.3"
 FT /product= "creatinase mutant Ctsd7"
 FT /note= "no stop codon given"
 XX
 XX EP1298213-A1.

PD 02-APR-2003.
 XX
 PF 17-SEP-2002; 2002EP-00020793.
 XX
 PR 20-SEP-2001; 2001EP-00121780.
 XX
 PA (HOPE) ROCHE DIAGNOSTICS GMBH.
 PA (HOPE) HOFMANN LA ROCHE & CO AG F.
 XX
 PI Shao Z, Schmuck R, Kratzsch P, Kenkies J, Weisser H;
 XX
 DR WPI; 2003-383834/37.
 DR P-PSDB; ABR43480.
 XX
 PT New variant of an Erwinia-type creatinase modified relative to a wild-
 PT type creatinase having creatinase activity, useful for determining
 PT creatinase and/or creatine concentration in a sample.
 PS
 XX Example 4; Page 40-42; 51pp; English.
 CC The present invention describes a variant of an Erwinia-type creatinase
 CC (I) modified relative to a wild-type creatinase having creatinase
 CC activity. The variant comprises at least one amino acid substitution at a
 CC position of the fully defined 404 amino acid sequence given in ABR43467;
 CC these are selected from N130, M203, I278, I1304 and F395. Creatinase has
 CC the EC number EC 3.5.3.3 and is also known as creatine amidohydrolase.
 CC Also described is a reagent (II) for determining creatine concentration in a
 CC sample.
 CC Erwinia-type creatinase variant. The variant is useful for determining
 CC creatinase and/or creatine concentration in a sample. Measuring
 CC creatinase and/or creatine are useful for diagnosing uraemia, chronic
 CC nephritis, acute nephritis, gigantism, tonic muscular dystrophy and some
 CC other related diseases. The mutant enzymes have improved stability, lower
 CC conductivity and/or lower Km-values for creatine; they are much better
 CC suited to detection methods for creatine. The present sequence encodes a
 CC mutant Erwinia creatinase from the present invention
 CC
 SQ Sequence 1212 BP; 239 A; 391 C; 362 G; 220 T; 0 U; 0 Other;
 Query Match 80.7%; Score 978.4; DB 8; Length 1212;
 Best Local Similarity 88.0%; Pred. No. 8.2e-157;
 Matches 1066; Conservative 0; Mismatches 146; Indels 0; Gaps 0;

QY 481 TCGCTGGAAGAGCAGAGCTGATCCGCGAAGGCGCCGCGTGTGTGACGTGCGCGCGG 540
 DB 481 TCGCTGGAAGAGCAGAGCTGATCCGCGAAGGCGCCGCGTGTGTGACGTGCGCGCGG 540
 QY 541 GCTTCGCGCGCTGCCATCAAGGCGCGCGTCCCGAGCATGAGTGGCCATTCGCAACACC 600
 DB 541 GCTTCGCGCGCTGCCATCAAGGCGCGCGTCCCGAGCATGAGTGGCCATTCGCAACACC 600
 QY 541 GCTTCGCGCGCTGCCATCAAGGCGCGCGTCCCGAGCATGAGTGGCCATTCGCAACACC 600
 DB 541 GCTTCGCGCGCTGCCATCAAGGCGCGCGTCCCGAGCATGAGTGGCCATTCGCAACACC 600
 QY 601 AATGGAGATCCGCGAGATGCGCAATTCGTTCCCTTCGTGAGGCTGATGACCACTTGG 660
 DB 601 AATGGAGATCCGCGAGATGCGCAATTCGTTCCCTTCGTGAGGCTGATGACCACTTGG 660
 QY 661 ACTGCTTCAGTCCGCGCATCAACACGCGCGCGCAACAATCCGCTCAACACCGCATC 720
 DB 661 ACTGCTTCAGTCCGCGCATCAACACGCGCGCGCAACAATCCGCTCAACACCGCATC 720
 QY 721 GTGCAATCCGCGCATCTCTTTCGCTCAACACTTCCCGATGATCTTCCGCTACTACAC 780
 DB 721 GTGCAATCCGCGCATCTCTTTCGCTCAACACTTCCCGATGATCTTCCGCTACTACAC 780
 QY 781 GCGCTGAGCGCAGCTGTTGCGAGCATGTCAGTGAACGCGAGCTGACATCTGGAG 840
 DB 781 GCGCTGAGCGCAGCTGTTGCGAGCATGTCAGTGAACGCGAGCTGACATCTGGAG 840
 QY 841 AAGAACGTGCGCGTGCATCGCGCGCGCTCGAGCTGATCAACCGCGCGCGCTGCAAG 900
 DB 841 AAGAACGTGCGCGTGCATCGCGCGCGCTCGAGCTGATCAACCGCGCGCGCTGCAAG 900
 QY 901 GACATCGCCATGACCTCAACGAGTGTACCGCGAGTGGAGACTGCTGAAGTACCGCTCC 960
 DB 901 GATATCGCCATGACCTCAACGAGTGTACCGCGAGTGGAGACTGCTGAAGTACCGCTCC 960
 QY 961 TTCGCTATGCGCACTCCCTGCGCGCTGCGTGCCTACTACAGTGGCGCGCGCGCTG 1020
 DB 961 TTCGCTATGCGCACTCCCTGCGCGCTGCGTGCCTACTACAGTGGCGCGCGCGCTG 1020
 QY 1021 GAGCTGCGCGAGGATCGACACCGAGCTGAAGCCCGCGCATGCTGCTCCATGAGCGCG 1080
 DB 1021 GAGCTGCGCGAGGATCGACACCGAGCTGAAGCCCGCGCATGCTGCTCCATGAGCGCG 1080
 QY 1081 ATGTGATGCTGCGCGAGGCGCATGCGCGCGCTGCGCGCTGATCGAGCAGCATCTCTG 1140
 DB 1081 ATGTGATGCTGCGCGAGGCGCATGCGCGCGCTGCGCGCTGATCGAGCAGCATCTCTG 1140
 QY 1141 ATGCTGCGCGAGGAGAGTGTGCGAGAACATCAACCGCGCTTCCGTTGCTCGGAGCAAC 1200
 DB 1141 ATGCTGCGCGAGGAGAGTGTGCGAGAACATCAACCGCGCTTCCGTTGCTCGGAGCAAC 1200
 QY 1201 ATCATCCGCAAC 1212
 DB 1201 ATCATCCGCAAC 1212

RESULT 15
 ACC69520
 ID ACC69520 standard; DNA; 1212 BP.
 XX
 AC ACC69520;
 XX
 DT 21-JUL-2003 (first entry)
 XX
 DE Mutant Erwinia creatinase CTs62 encoding DNA SEQ ID NO:19.
 XX
 KW Erwinia; creatinase; creatine amidohydrolase; enzyme; EC 3.5.3.3;
 KW Erwinia-type creatinase; creatine; creatinine; uraemia; gigantism;
 KW chronic nephritis; acute nephritis; tonic muscular dystrophy; mutant;
 KW gene; ds.
 XX
 OS Erwinia sp.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 7, 2005, 20:16:05 ; Search time 4424 Seconds
(without alignments)
10428.096 Million cell updates/sec

Title: US-10-807-228a-2

Perfect score: 1212

Sequence: 1 ATGACAGCAGCATCTTGCA.....AACACACATCATCCGCAAC 1212

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_ests1:*
2: gb_ests2:*
3: gb_hrc:*
4: gb_ests3:*
5: gb_ests4:*
6: gb_ests5:*
7: gb_ests6:*
8: gb_gsa1:*
9: gb_gsa2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	322.6	26.6	623	4	BI376242 BFLG3 000
2	221	18.2	470	1	AI906433 IL-BRT09-
3	207.8	17.1	494	6	CD307119 StrPu691.
4	186.6	15.6	442	6	CD295681 StrPu691.
5	182.6	15.1	821	7	CV214674 EST874384
6	135	11.1	670	7	CV214629 EST874339
7	122	10.1	519	6	CD296311 StrPu691.
8	83.2	6.9	551	7	CP919042 BFL0531.
9	83.2	6.9	553	4	BI387857 BFL26_002
10	73.8	6.1	2529	9	CL964961 OsIFCC011
11	69.6	5.7	524	8	BZ894814 Hg4 0102
12	69.2	5.7	722	8	CB656525 OSUNEcl0P
13	67.6	5.6	722	6	CB683938 OSUNEcl0P
14	66	5.4	2538	9	CL963721 OsIFCC038
15	65	5.4	1368	9	CL948585 OsIFSB000
16	65	5.4	1368	9	CL948585 OsIFSB000
17	63.6	5.2	925	9	CNS0091P
18	61.6	5.1	1115	7	CK208301 FGAS02000
19	61.4	5.1	707	7	CF874011 trlC035Xc
20	61.4	5.1	732	7	CF886540 trlC030Xc
21	61.4	5.1	763	6	CB903555 trlC035Xc
22	61.4	5.1	802	6	CB902534 trlC030Xc
23	60.8	5.0	712	6	CB869465 AZ02.111L
24	60.8	5.0	812	6	CB659172 OSUNEcl150

25	60.6	5.0	604	1	AU162766	AU162766
26	60.4	5.0	1575	9	CL979927	CL979927 OsIFCC045
27	60.2	5.0	1116	9	CL963961	CL963961 OsIFCC010
28	60.2	5.0	2022	9	CL973018	CL973018 OsIFCC023
29	59.8	4.9	635	6	CA254345	CA254345 SCBPR411
30	59.8	4.9	748	6	CA227320	CA227320 SCULF301
31	59.8	4.9	1602	9	CL975927	CL975927 OsIFCC043
32	59.6	4.9	993	9	CL977070	CL977070 OsIFCC044
33	59.4	4.9	655	6	CA200834	CA200834 SCRFL107H
34	59	4.9	579	2	BE040798	BE040798 OPIG11 O
35	59	4.9	668	6	CD225047	CD225047 CCC1.37 H
36	59	4.9	690	2	BE041110	BE041110 OPI5H01 O
37	58.8	4.9	512	4	BM140351	BM140351 WHE0474_a
38	58.8	4.9	614	2	BF202900	BF202900 WHE1782 H
39	58.8	4.9	888	9	CG440030	CG440030 CGVHP107H
40	58.6	4.8	719	7	CF487085	CF487085 POL1.41 D
41	58.6	4.8	937	7	CG240162	CG240162 OCYAE68TV
42	58.4	4.8	748	7	CN143339	CN143339 WOUNDI_15
43	58.4	4.8	768	7	CN148454	CN148454 WOUNDI_56
44	58.4	4.8	814	7	CN126847	CN126847 RHOH1_19
45	58.2	4.8	1056	7	CK212443	CK212443 FGAS02431

ALIGNMENTS

RESULT 1
BI376242
LOCUS
DEFINITION
BFLG3_000038 Amphioxus 5-6 hrs cDNA library (Name convention: BFLG or MPMGP498) Branchiostoma floridae cDNA clone MPMGP498A1518 5', mRNA sequence.

ACCESSION
BI376242
VERSION
BI376242.1 GI:30911206
KEYWORDS
EST.
SOURCE
Branchiostoma floridae (Florida lancelet)
ORGANISM
Branchiostoma floridae
Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae; Branchiostoma.

REFERENCE
1 (bases 1 to 623)
Panopoulou, G., Hennig, S., Groth, D., Krause, A., Pousetka, A., J., Herwig, R., Vingron, M. and Lehrach, H.
New evidence for genome-wide duplications at the origin of vertebrates using an amphioxus gene set and completed animal genomes

JOURNAL
Genome Res. 13 (6A), 1056-1066 (2003)
MEDLINE
22683279
PUBMED
12799346

COMMENT
Contact: Panopoulou G
Laboratory 145, dept. Lehrach
Max-Planck-Institut fuer Molekulare Genetik
Inestr.63-73, D-14195 Berlin, Germany
Tel: +49 30 8413 1235
Fax: +49 30 8413 1128
Email: panopoulou@molgen.mpg.de
The library was characterised by oligonucleotide fingerprinting (ONFP) to reduce sequencing redundancy. According to the ONFP procedure, clones giving the same hybridisation pattern with a battery of 200 Bmer oligonucleotides are grouped into clusters. One clone per cluster is selected for sequencing. The size of each cluster is an indicator of the frequency of a transcrip in the analysed library. The cluster size as well the coordinates of the rest of the clones assigned to the same fingerprint cluster as the clone from which the above EST is generated is available at the amphioxus project site at: <http://www.molgen.mpg.de/amphioxus/>
Clones and filters are distributed via the Resource Center/Primary Database of the German Genome Project (<http://www.rzpd.de>)
PCR primers
FORWARD: 5' CCCGAGCTTACCTATGATCCGCGCTG 3' (M1385P)
BACKWARD: 5' GCTATTACGCGAGCTGGGAAAGGGGATG 3' (M1385P)
Insert Length: 1200 Std Error: 0.00
Seq primer: 5'-CCGTCGCGAATTCGCGGT-3' pSPORT3/86
High quality sequence stop: 623.

FEATURES
source

Location/Qualifiers
1..623
/organism="Branchiostoma floridae"
/mol_type="mRNA"
/db_xref="taxon:7739"
/clone="MPMGp498A1518"
/cissue_type="whole embryo"
/dev_stage="5-6 hrs (gastrula stage)"
/lab_host="E.coli, XL1 blue"
/clone_1ib="Ampliphox 5-6 hrs cDNA library (Name convention: BFLG or MPMGP498)"
/note="Vector: pSPORT1; Site 1: SalI; KpnI, EcoRI (5'); Site 2: NotI, BamHI, HindIII (3'); OligodT primed and directionally cloned in pSPORT1 vector using a NotI (5'-PACTAGTCTGATCGGAGCGCGGCC (T)15-3' and a SalI 5'-TCGACCCAGCGCTCG-3' adapters (Gibco BRL)."

ORIGIN

Query Match 26.6%; Score 322.6; DB 4; Length 623;
Best Local Similarity 70.4%; Pred. No. 1.3e-60;
Matches 430; Conservative 0; Mismatches 181; Indels 0; Gaps 0;

570 GCCCGAGCATGAGTGGCCATGCCACCAATGCGATGATCCGAGATGCCAATC 629
13 GCTGATGATGAGTGGCGATGACGCCCTCCAAACAATGATACGGAAATTGGCAGAC 72
630 GTTCCCTTCTGTGAGCTGATGACACTGTGACCTGTTCAGTCCGGCATCAACCGA 689
73 CTACCGCATGAGGGAATCATGAACTTGGCGCTGATTCAGTCAAGGATCAACACTGA 132
690 CGGCGCGCAATCCGCTGACCAACCGCATGCGCAATCCGCGCAATCCTTCCGTCA 749
133 CGGGGACCAACCTCTGACGTCAAGCGCGCGGCGAGTGGGAACTCTGTCACTCA 192
750 CACTTCCCGATGATCTTGGCTACTACACCGCGCTGAGAGCGCACGCTGTTCGACCA 809
193 CACTTCCGAAATGATGGCGGATCACTATGCGATGAGAGGAAACGCAATTCGTAACA 252
810 TGTGATGAGCGCAAGCTGACATCTGGAGAAAGAACTGGCCGTGATCGCGCGGCT 869
253 CGCCTCCGACGACACTCCGACTTGGGAGATCAACTTAAGTAACTCGCGCGGCGCT 312
870 CGAGCTGATCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 929
313 GGAAGCTGATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 372
930 CGCGAGTGGAGCTGCTGAAGTACCGCTCTTCCGCTATGCGCACTCTTCCGCGTCT 989
373 CGGGAAACAGGGCTGCTGATGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 432
990 GTGCACTAATAAGTCTGCGAGGCGCGCGCTGAGCTGCGCGAGGACATGCAACCGAGCT 1049
433 GTGCACTAATAAGTCTGCGAGGAGAGATTTGGAGCTGAGAGAGAGAGAGAGAGAGAG 492
1050 GAAAGCGCGGATGCTGCTGATGAGAGCGGATGAGTGTGCTGCGGAGGAGAGAGAGAG 1109
493 GGAAGCGCGGATGCTGCTGATGAGAGCGGATGAGTGTGCTGCGGAGGAGAGAGAGAG 552
1110 TCCCGGCGGCTATGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1169
553 AGCGGCGGCTATGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 612
1170 CACCGGCTTCC 1180
613 TACAGGCTCC 623

RESULT 2
A1906433/c 470 bp mRNA linear EST 30-MAR-2000
LOCUS A1906433
DEFINITION IL-BT109-280199-002 BT109 Homo sapiens cDNA, mRNA sequence.
ACCESSION A1906433
VERSION A1906433.1 GI:6496820

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT

EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 470)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Mateukuma, A., Bais, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
10737800
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the PAPESP/ILICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/seq/gethtml.pl?tl=IL&t2=IL-BT109-002.html
&t3=280199&t4=1)
Seq primer: puc 18 forward.
Location/Qualifiers
1..470
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/sex="female"
/dev_stage="Adult"
/clone_1ib="BT109"
/note="Organ: breast; Vector: puc18; Site 1: SmaI; Site 2:
SmaI; A mini-library was made by cloning products derived
from ONESTES PCR (U.S. Letters Patent application No.
196,716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."

ORIGIN

Query Match 18.2%; Score 221; DB 1; Length 470;
Best Local Similarity 69.4%; Pred. No. 3.1e-38;
Matches 313; Conservative 0; Mismatches 137; Indels 1; Gaps 1;

630 GTTCCCTTCTGTGAGCTGATGACACTGACCTGTGTTCAATGCGGATCAACCGA 689
455 GATCCCTGAATGCGGCTGATTAACAACCTTGACATGTTCCAGTCCGGGATCAACCGA 396
690 CGGCGCGCAATCCGCTGACCAACCGGATGAGTGAATCCGGCGCATCTTCCGTCA 749
395 TGGGGCGATTAACCGGCTGACCAACCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 336
750 CACTTCCCGATGATCTTGGC-TACTACACGCGCTGAGAGCGACGCTGTTCTGCGAC 808
335 TTGCTTCCCGATGATGCGCGGCAATTAACAACGCGCTTGAAGGACGCGTTCCTGACCC 276
809 ATGTGATGACCGCACCTGACATCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 868
275 ACTGCCCGAGAGAGTACCTACGCTGTGCGAGGCGCAAGTGAAGTGAAGAGAGAGAG 216
869 TCGAGCTGATCAAGCGCGGCGCGCGCGCTGCAAGAGAGAGAGAGAGAGAGAGAGAGAG 928
215 TGAACCTGAGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 156
929 ACCGCGAGTGGAGACTGCTGAAGTACCGCTCTTCCGCTATGCGCACTCTTCCGCGTGC 988

Db 155 TCTCGCGCAGACGCTGCTGACGACCTTCGCGTCTGTCGCGACCT 96

Qy 989 TGTGCACTACTAGTGTGCGAGGCGCGTGTGAGTGTGCGAGACATGACCGAGC 1048

Db 95 TGAGCCACTACTATGCGCCGAAAGCGGCGTGTGAGCTGTGCGAGACATGACCGGTGC 36

Qy 1049 TGAAGCCCGGAGTGTGTCTTCATGAGGCC 1079

Db 35 TGAACCGGCGATGTGTGTCTCATGACCGCC 5

RESULT 3
CD307119/c 494 bp mRNA linear EST 16-SEP-2003
LOCUS Strongylocentrotus purpuratus cDNA clone
DEFINITION MPMG691E0990/MP1_SURUDI_90B9 5', mRNA sequence.

ACCESSION CD307119 GI:34752168
VERSION CD307119
KEYWORDS EST.
SOURCE Strongylocentrotus purpuratus
ORGANISM Strongylocentrotus purpuratus
Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
Echinoidea; Euechinoidea; Echinacea; Echinoida;
Strongylocentrotidae; Strongylocentrotus.
1 (bases 1 to 494)

REFERENCE Pouska, A.J., Groth, D., Hennig, S., Thamm, S., Cameron, A., Beck, A.,
Reinhardt, R., Herwig, R., Panopoulou, G. and Lehrach, H.
Generation, annotation, evolutionary analysis, and database
integration of 20,000 unique sea urchin EST clusters
Genome Res. 13 (12), 2736-2746 (2003)

AUTHORS Pouska, A.J.
TITLE Laboratory 145, dept. Lehrach
JOURNAL Max-Planck-Institut fuer Molekulare Genetik
COMMENT Insestr.63-73, D-14195 Berlin, Germany
Tel: +49 30 8413 1235
Fax: +49 30 8413 1128
Email: pouska@molgen.mpg.de

FEATURES
source
1.494
Location/Qualifiers
/organism="Strongylocentrotus purpuratus"
/mol_type="mRNA"
/db_xref="taxon:7668"
/clone="MPMG691E0990/MP1_SURUDI_90B9"
/tissue_type="whole larva"
/dev_stage="larva 2-3 weeks"
/lab_host="E.coli, XL1 blue"
/note="Vector: pSport1; Site 1: NotI; Site 2: SalI; Random
primed and directionally cloned in pSport1 vector using a
NotI (5'-TGACTAGTCTAGTGGCGAGCGCGCC (7)15-3' and a
SalI 5'-TCGACCCACGCGTCCG-3' adapters (Gibco BRL)"

High quality sequence stop: 494.

PCR Primers
FORWARD: 5' CCCGAGCTTTACCTTATGCTCCGCGCTCG 3' (M13RSP) 5'-seq
BACKWARD: 5' GCTATTAGCCGAGTGGGAGAGGAGATGTG 3' (M13FSP) 3'-seq
Seq primer: 5'-CCGCTCCGGAATCCGCGGT-3' pSport1/86

ORIGIN
Query Match 17.1%; Score 207.8; DB 6; Length 494;
Best Local Similarity 66.8%; Pred. No. 2.5e-35;

Matches 312; Conservative 0; Mismatches 152; Indels 3; Gaps 1;

Qy 730 GGCACATCCTTTGCTGCTCAACACCTTCCGATGATCTGGCTACTACACCGGCTGAG 789

Db 476 GGTAGGCGCATGCTCTTAAAGCTTCTTATATAGCCGCTTATTTCTGGGCTTGA 417

Qy 790 CGCAGCTGTTCTGCAACCATGTGATGACCGCAGC---CTGACATCTGGGAAAGAC 846

Db 416 CGAACCTTTTCTTGAACACAGTGCCTTCCGACCGGCATCTTAAAGTGGCAGGTCA 357

Qy 847 GTGGCCGTGATGCGCGCGGCTCGAGCTGATTAAGCCGGCGCGCTGCAAGACATC 906

Db 356 TGCAGATTCACCGCGCGGAGTGAAGCTCATCAACCTGTGTCAAACTCTCGACGTC 297

Qy 907 GCCATGAGCTCAAGAGATGTAAGCGGAGTGGGACCTGTGAAGTACCGCTCTTCGGC 966

Db 296 GCGCTGAGTTGAACAGATGTAACCGGAGAAACCTTACATGACAAAGACTTCGGC 237

Qy 967 TATGCGCACTCTTCCGCGTGTGCTGCACTACTACGCTCGCAGGCGCGCTGAGCTG 1026

Db 236 TAGGGCATCTCTTGGGGTCTATGCGACTACTACGCGCGGAGCAGCTCTTAGTTG 177

Qy 1027 CCGAGGACATGACACCGAGCTGAAGCCCGCATGTGTCTTCAATGAGCCGATGTC 1086

Db 176 AGGAGACATGACGAGCTGTATACAGCTGTATGTGTCTTCCATGAGCCACATCTC 117

Qy 1087 ATGCTGCGGAGGCGATGCGCGGTGCGCGGCTATCGGAGACACATCTGATGTC 1146

Db 116 ACATATCGACAGACAGACGCTGTGAGAGGCTACAGGAGATATATCATGTTGTG 57

Qy 1147 GGGGAGACGCTGCCGAGACATCAACGCGCTTCCCGTGGTCCGGA 1193

Db 56 ACTGAAGAGGCGCGCTGCTTCAATACCGCATTTCTTAGGACCGGA 10

RESULT 4
CD295681 442 bp mRNA linear EST 16-SEP-2003
LOCUS Strongylocentrotus purpuratus cDNA clone
DEFINITION MPMG691E0990/MP1_SURUDI_90B9 3', mRNA sequence.

ACCESSION CD295681
VERSION CD295681.1 GI:34746758
KEYWORDS EST.
SOURCE Strongylocentrotus purpuratus
ORGANISM Strongylocentrotus purpuratus
Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
Echinoidea; Euechinoidea; Echinacea; Echinoida;
Strongylocentrotidae; Strongylocentrotus.
1 (bases 1 to 442)

REFERENCE Pouska, A.J., Groth, D., Hennig, S., Thamm, S., Cameron, A., Beck, A.,
Reinhardt, R., Herwig, R., Panopoulou, G. and Lehrach, H.
Generation, annotation, evolutionary analysis, and database
integration of 20,000 unique sea urchin EST clusters
Genome Res. 13 (12), 2736-2746 (2003)

AUTHORS Pouska, A.J.
TITLE Laboratory 145, dept. Lehrach
JOURNAL Max-Planck-Institut fuer Molekulare Genetik
COMMENT Insestr.63-73, D-14195 Berlin, Germany
Tel: +49 30 8413 1235
Fax: +49 30 8413 1128
Email: pouska@molgen.mpg.de

The library was characterized by oligonucleotide fingerprinting (ONF) to reduce sequencing redundancy. According to the ONF procedure, clones that display the same hybridisation matrix with a battery of 200 Bmer oligonucleotides are grouped into clusters. One clone per ONF cluster is selected for sequencing. The size of each cluster is an indicator of the frequency of a transcript in the analysed library. The cluster size as well as the coordinates of the other clones assigned to the same ONF cluster as the clone from which the above EST is generated is available at the sea urchin project web site at: http://www.molgen.mpg.de/ag_seaurchin/. cDNA clones and filters are distributed via the Resource Center/Primary

Database of the German Human Genome Project (<http://www.rzpd.de>)
 PCR PRIMERS
 FORWARD: 5' CCCAGGCTTTACATTTATGCTCCGGCTCG 3' (M13RSP) 5'-seq
 BACKWARD: 5' GCTATTACGCGAGCTGGCAAGGGGAGATG 3' (M13FSP) 3'-seq
 Seq primer: 5' GCTATTACGCGAGCTGGCAAGGGGAGATG 3' (M13FSP)
 High quality sequence stop: 442.

FEATURES

source

1..442
 /organism="Strongyloides purpuratus"
 /mol_type="mRNA"
 /db_xref="taxon:7668"
 /clone="MPMG691B0990.MPI SURUDI_90B9"
 /issue_type="whole larva"
 /dev_stage="larva 2-3 weeks"
 /lab_host="E.coli, XL1 blue"
 /clone_lib="Sea urchin larva cDNA library MPMG691"
 /note="Vector: pSport1; Site 1: NotI; Site 2: SalI; Random
 primed and directionally cloned in pSport1 vector using a
 NotI (5'-pGACTAGTCTAGATCGCGGCGGCC (r)15-3' and a
 SalI 5'-TCGACCCACGCGTCCG-3' adapters (Gibco BRL)"

ORIGIN

Query Match 15.6%; Score 188.6; DB 6; Length 442;
 Best Local Similarity 67.3%; Pred. No. 4.2e-31;
 Matches 282; Conservative 0; Mismatches 134; Indels 3; Gaps 1;

QY 730 GGGGACATCTTGGCTCAACACCTTCCGATGATCTTGGCTACTACACCGCGCTGGAG 789
 DB 20 GGTGACGCGATCGTCTTAACGCTTCCCTATGATAGCGGCTATTCTCGGCTTGAA 79
 QY 790 CGACGCTTTCGCGACCAATCGATGACGCGAGC---CTGACATCGTGGGAGAGAAC 846
 DB 80 CGAACCTTTCTTGAACCAAGCTTCCGACGCGCATTTGAAGTGGCGATCAC 139
 QY 847 GTGCGCTGATCGCGCGGCTCGAGCTGATCAACCGCGCGCTGACAGACATC 906
 DB 140 TCGACGTTTCAACGCGCGGAGATGAGCTCATCAAGCTTGATCAATCTCGACGTC 199
 QY 907 GCGATGAGCTCAACAGATGATCCGCGAGTGGAGCTTGTAAGTACCGCTCTTGGC 966
 DB 200 GGGCTGAGTTGAAACAGATGATCCGCGAGAGAACTTACTACATCAAGAGCTTCGCG 259
 QY 967 TATGCGCATCTCTTCCGCGCTGCTGCTGCTACTACTATGATGCTCGAGGCGCGCTGAGAC 1026
 DB 260 TACGGGCACTCTTCCGCGCTGCTGCTGCTACTACTACTGCGCGCGAGAGCTTGAAGT 319
 QY 1027 CGGAGAGCATGACACCGAGCTGAAGCCGCGAGTGGTCTCCATGAGCCGATGATG 1086
 DB 320 AGGAGAGCATGACAGCTGTTATACAGCTGATGCTGCTTCCATGAGCCACATCTC 379
 QY 1087 ATGCTCCGAGAGGAGATGCGCGGCTGCGCGCTATGCGAGACAGCATCTGATCT 1145
 DB 380 AGATATACAGAGAGAGAGCTGCTGAGAGAGGCTACAGGAGCATATATCATGGGTGT 438

RESULT 5

CV214674

821 bp mRNA linear EST 16-SEP-2004

LOCUS CV214674 non-normalized T1 cDNA library Trichomonas vaginalis cDNA
 DEFINITION clone T1TC244 5' end, mRNA sequence.

ACCESSION CV214674
 VERSION CV214674.1 GI:52161654
 KEYWORDS EST.
 SOURCE Trichomonas vaginalis
 ORGANISM Trichomonas vaginalis
 Eukaryota; Parabasalida; Trichomonadinae; Trichomonas.

REFERENCES

1. (bases 1 to 821)
 Carlton, J.M., Dyall, S., Johnson, P.J. and Fraser, C.M.
 The complete genome sequence of the sexually transmitted parasite
 Trichomonas vaginalis
 Unpublished (2004)
 CONTACT: Jane Carlton

Parasite Genomics Group
 The Institute for Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-530-9319
 Fax: 301-838-0208
 Email: carlton@tigr.org
 Seq primer: lambda Triplex2.

FEATURES

source

1..821
 /organism="Trichomonas vaginalis"
 /mol_type="mRNA"
 /strain="T1"
 /db_xref="taxon:5722"
 /clone="T1TC244"
 /clone_lib="non-normalized T1 cDNA library"
 /note="Vector: Lambda Triplex2; Site 1: SfiI; Site 2:
 SfiI; T. vaginalis strain T1 library constructed from
 cDNA, made in lambda Triplex2. Inserts cloned
 unidirectionally in the SfiI and SfiI sites. Mase excision
 of library produced inserts in pTriplex2 plasmid. Inserts
 sequenced from both 5' and 3' ends using Triplex2
 sequencing primer and polydT 24 bp primer respectively."

ORIGIN

Query Match 15.1%; Score 182.6; DB 7; Length 821;
 Best Local Similarity 60.2%; Pred. No. 9.2e-30;
 Matches 327; Conservative 0; Mismatches 204; Indels 12; Gaps 1;

QY 35 ACAACGCGGAGAAAGATTATTCGCGTTTTCGATGCGAGATGACCGCGCCAAACG 94
 DB 279 ACAACGCGGAGAAAGATTATTCGCGCTGACCTTCTCAGACCGAGTTCGAGCGCTCTCCG 338
 QY 95 AGCTTCGCGCGTGAATGCGCAAGAACATGTCGATGCGCGCTGTACCTTTACCT 154
 DB 339 GCGTCGCGCGATCATGCGCGGAGAAAGCTTGAAGCGCTCATCTCACACGATCACT 398
 QY 155 GCATCAATCACTATTCCGCGCTGCTGTACTGTCTATTTCGAGCAAGTACGCGATGTA 214
 DB 399 CATCAAGTACTACTTCCACTTCTTCACTTCTTGCGCGCTCTTACGCGATGATG 458
 QY 215 TCGACCAACAACGACGACGAGATTTTCGCGCGGATCGAGCGCGGCGGCGGCGG 274
 DB 459 TCACCAAGATGACACGCGTCACTATCAACGCGCAATATGACCGCGGAGATGCTTGGCG 518
 QY 275 GCACTTCGCGCAACAATCACTACCTACCTGCGCGCGGCAATTTCTATCGGCG-- 332
 DB 519 GCACTTCGCGCAAGAACTGTGTACAGGACTGCGCGGAGATTAATCAATCAACGCA 578
 QY 333 -----CGTGGCGCGAGTGAACAAGCGCGCAAGCGCATCGCATCGAGTTGAC 382
 DB 579 TCGAGGAAGTCTTGGCACCGCGGATATCAACCGCGCGCATTCGCGCGATGATG 638
 QY 383 AGCTCAATCTGACTTTCGCGCGCGGCTGAGAGAGCCCTTACCGCGCGTGAAGTTG 442
 DB 639 CGTGGCGCGTGAACAACGCAACAGATCAAGATCAAGCGCGCTTCTCGGAGGAGAG 698
 QY 443 ACATCAAGCAAGCGCTTCGATGATGATGCGGACCATCAAGTGCCTCGAAGAGCAAG 502
 DB 699 AGTTCGCGGACGAGCGATGCGCGAGCATGATCAAGTCCCGCGAGAGATGCGGCTCA 758
 QY 503 TCGGAGAAAGCGCGCGGTGTGACGTCGCGCGCGCGCTGCGCGGCTGCATCAAG 562
 DB 759 TCAAGACGATGCGCGCATTTGATCTTGGTGGTGGAGGCGCATCGCAACGCGCAT 818
 QY 563 CCG 565
 DB 819 CCG 821

RESULT 6

CV214629

670 bp mRNA linear EST 16-SEP-2004

LOCUS CV214629 non-normalized T1 cDNA library Trichomonas vaginalis cDNA
 DEFINITION EST874339

accession clone TVTC131 5' end, mRNA sequence.
 version CV214629
 key words EST
 source Trichomonas vaginalis
 organism Trichomonas vaginalis
 reference Eukaryota; Parabasalida; Trichomonada; Trichomonadida;
 Trichomonadidae; Trichomonadinae; Trichomonas.
 authors 1 (bases 1 to 670)
 title Carlton, J.M., Dyall, S., Johnson, P.J. and Fraser, C.M.
 journal The complete genome sequence of the sexually transmitted parasite
 Trichomonas vaginalis
 comment Unpublished (2004)
 contact: Jane Carlton
 Parasite Genomics Group
 The Institute for Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850, USA
 tel: 301-530-9319
 fax: 301-838-0208
 email: carlton@tigr.org
 seq primer: lambda Triplex2.
 features
 source
 location/Qualifiers
 1..670
 /organism="Trichomonas vaginalis"
 /mol_type="mRNA"
 /strain="T1"
 /db_xref="taxon:5722"
 /clone="TVTC131"
 /note="lib=non-normalized T1 cDNA library"
 /note="Vector: Lambda Triplex2; Site 1: SfiA; Site 2:
 SfiB; T. vaginalis strain T1 library constructed from
 cDNA, made in lambda Triplex2. Inserts cloned
 unidirectionally in the SfiA and SfiB sites. Mass excision
 of library produced inserts in Triplex2 plasmid. Inserts
 sequenced from both 5' and 3' ends using Triplex2
 sequencing primer and polydT 24 bp primer respectively."
 origin
 Query Match 11.1%; Score 135; DB 7; Length 670;
 Best Local Similarity 65.3%; Pred. No. 2.8e-19;
 Matches 198; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

accession CD296311
 version CD296311.1 GI:34747388
 key words EST.
 source Strongyloides purpuratus
 organism Strongyloides purpuratus
 reference Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
 Echinoida; Echinoida; Echinodermata; Echinodermata;
 Strongyloides; Strongyloides.
 authors 1 (bases 1 to 519)
 title Reinhardt, R., Herwig, R., Panopoulou, G. and Lehnach, H.
 journal Generation, annotation, evolutionary analysis, and database
 integration of 20,000 unique sea urchin EST clusters
 genome Res. 13 (12), 2736-2746 (2003)
 comment Contact: Pousetka, A.J.
 laboratory 145, dept. Lehnach
 Max-Planck-Institut fuer Molekulare Genetik
 Ihnestr.63-73, D-14195 Berlin, Germany
 tel: +49 30 8413 1235
 fax: +49 30 8413 1128
 email: pousetka@molgen.mpg.de
 The library was characterised by oligonucleotide fingerprinting
 (ONP) to reduce sequencing redundancy. According to the ONP
 procedure, clones that display the same hybridisation matrix with a
 battery of 200 8mer oligonucleotides are grouped into clusters. One
 clone per ONP cluster is selected for sequencing. The size of each
 cluster is an indicator of the frequency of a transcript in the
 analysed library. The cluster size as well as the coordinates of
 the other clones assigned to the same ONP cluster as the clone from
 which the above EST is generated is available at the sea urchin
 project web site at: <http://www.molgen.mpg.de/seq/seaurchin/>. cDNA
 clones and filters are distributed via the Resource Center/Primary
 Database of the German Human Genome Project (<http://www.rzpd.de>)
 PCR Primers
 FORWARD: 5' CCCAGCGCTTACCTTATGCTCCGGCTCG 3' (M13RSP) 5'-seq
 BACKWARD: 5' GCTATGACCGCAGCTGGGAGGAGATGTC 3' (M13RSP) 3'-seq
 Seq primer: 5'-CCGTCGCGAATTCGCCGGCT-3' pSPORT3/86
 High quality sequence stop: 519.
 features
 source
 location/Qualifiers
 1..519
 /organism="Strongyloides purpuratus"
 /mol_type="mRNA"
 /db_xref="taxon:7668"
 /clone="MPMG691N125;MPI SURUDI_15N12"
 /issue_type="whole larva"
 /dev_stage="larva 2-3 weeks"
 /lab_host="E.coli, XL1 blue"
 /clone_lib="Sea urchin larva cDNA library MPMG691"
 /note="Vector: pSPORT1; Site 1: NotI; Site 2: SalI; Random
 primed and directionally cloned in pSPORT1 vector using a
 NotI (5'-pACTTACTTATGATGCGAGCGGCC (7)15-3' and a
 SalI 5'-TCGACCCACGCTCG-3' adapters (Gibco BRL)."
 origin
 Query Match 10.1%; Score 122; DB 6; Length 519;
 Best Local Similarity 55.7%; Pred. No. 2e-16;
 Matches 274; Conservative 0; Mismatches 209; Indels 9; Gaps 2;

Oy	ATCGACGGCGGCAGACCCCTGGCG-----CCGCAGCTTGGCGAGCAACATCACTCAACACC	303
Oy	250	
Db	GTGCAGACAGCGGCAGAGCCTGGCGAGATCAACCCTGGAGAGCAAGATGTGGTCAATCAACAG	323
Db	264	
Oy	GACTGGCGCCCGGCAGCAATTTCTATGCGCCGCT---GCGCAGCTGAACACAGGCGCCCAAG	360
Oy	304	
Db	GATTGGCAATCGGCAACCTTCTGGAGCGCCGCTTGGCACTGTCTCGGCAACCTTTCGGAG	383
Db	324	
Oy	CGCATCGGCATCGAGTTGCAGCAGCATCAATTCGACTTCCGCGCGCAGCTCGAGGAAGCC	420
Oy	361	
Db	AAGATCGGAGCGCGAATTTGATCATATTACTCATGGAGAGCAACTAAAGTTAGAACAGGT	443
Db	384	
Oy	CTACCGGCGCTCGAGTTGTCGCAATCAAGCAGCCCTCGATGTGATGTCGCACATCAAG	480
Oy	421	
Db	GTAATTCAGACGACACCGTCGATGTCATATCCATCATCAATCGAATGCATATCAAG	503
Db	444	
Oy	TCGCTCGAAGAG	492
Oy	481	
Db	TCCTCCGAGAG	515

RESULT 8	LOCUS	DEFINITION
CE919042	CE919042	551 bp mRNA linear EST 05-NOV-2001
	BFL0551.000127	Amphioxus 26 hrs cDNA library (Name convention: BFL26 or MPMG531) Branchiostoma floridae cDNA clone MPMG531.07115; BFL26_115L7 5', mRNA sequence.

ACCESSION	CF919042
VERSION	CP919042.1
KEYWORDS	EST.
SOURCE	Branchiostoma floridae (Florida lancelet)
ORGANISM	Branchiostoma floridae

REFERENCE
AUTHORS
TITLE

1 (bases 1 to 551)
Panopoulou, G., Hennig, S., Groth, D., Krause, A., Poustka, A. J.,
Hewig, R., Vingron, M. and Lehrach, H.
New evidence for genome-wide duplications at the origin of
the vertebrate lineage: the *hox* and *eng* gene clusters and *eng* orthologs

JOURNAL
MEDLINE
PUBMED
COMMENT

Genome Res. 13 (6A), 1056-1066 (2003)
22683279
12799346
Contact: Panopoulos G
11015 2003-07-19

Email: panopoul@mpg.de
The library was characterized by oligonucleotide fingerprinting (ONFP) to reduce sequencing redundancy. According to the ONFP procedure clones giving the same hybridisation pattern with a

all clones described in this paper are available from which the above EST is generated is available at the project site at <http://www.molgen.mpg.de/amphioxus>. Clones and filters are distributed via the Resource Center/Primary Database of the German Genome Project (<http://www.izpdp.de>).

```

FEATURES
  source
    1. 551
      /organism="Branchiostoma floridae"
      /mol_type="mRNA"

```

```

/strain="wild type"
/db_xref="taxon:7739"
/clone="MMPG531L07115,BFL26_115U7"
/tissue_type="whole embryo"
/dev_stage="26 hrs (neurula stage)"
/lab_host="Escherichia coli, XLI blue"
/clone_lib="Amplioxus 26 hrs cDNA library (Name
convention: BFL26 or MMPG531)"
/notes="Vector: pSport1; Site_1: SalI, KpnI, EcoRI (5');
Site_2: NotI, BamHI, HindIII (3'); OligodT primed and
directionally cloned in pSport1 vector using a NotI
(5'-pGACTAGTCTAGATCGACGACGCGCGCC (7)15-3' and a SalI 5'-
TCGACCCACGCGTCCG-3' adapters (Gibco BRL)."
```

Query Match	6.9%	Score 83.2	DB 7	Length 551
Best Local Similarity	57.8%	Pred. No. 7.1e-08		
Matches 148	Conservative 0	Mismatches 108	Indels 0	Gaps 0
26 TGAATGGCAACAGCGCGAAGATATATCGCGCTTTTGAGATGCCAGATGACCCGCC 85				

Db 286 TGGCATGCAATACGGGCGAAGAGTCAACCCGACTTCTCAGTGGAGGAATCCAGAGAA 345
 Qy 86 GCCAAACGACCTTTCGCGCTGGATGGCCAAAGAACATGTGAGTGGCGGCTGTTCACT 145
 Db 346 GGGTGGACAAAGCTGCGTCTCATGTGTACGACGATATATAGATGACGCTCTTCACTT 405

Qy 146 CTTATCAGTCAGCACTACTATTCCGGCGGGCTGCTCTATTGGAGCGCAAGACG 205

Db 406 CCTATCATTAATCAACTACTACTGCGCATTTCTGTACATTCTGGTCCGACTACG 465

Qy 206 GCATGGTCATCGACCAACAAACGCCAGCGATTTCCGGCGGTCATGACGGCGGCCAGC 265

Db 466 GGCTGGTCTGTCACCATGAGCAAAAGTGTCTTCTTATCCACGCGGTGACGGTGGTCTCAGC 525

Oy 266 CCTGGCGCCGCGAGCTT 281
||||| | |||||

Db 526 CCTGGAGAAAGACAT 541

	EST	26-AUG-2003
RESULT 9		
B1387857		
LOCUS		
DEFINITION		
	553 bp	
B1387857	mRNA	linear
B1387857	Amphioxus 2chr cDNA library	(Name convention: BFL26)
B1387857	Branchiostoma floridae cDNA clone MPMG5331l07115 5'	MPMG5331

ACCESSION	BI387857	
VERSION	BI387857.1	GI:30922696
KEYWORDS	EST	
SOURCE	Branchiostoma floridae (Florida lancelet)	
ORGANISM	Branchiostoma floridae	
	Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae; Branchiostoma	
REFERENCE	Branchiostoma (t. 553)	

JOURNAL	Genome Res. 13 (6A), 1056-1066 (2003)
MEDLINE	22683279
PUBMED	12799346
COMMENT	Contact: Panopoulos G. - gpanopoulos@ucla.edu

The library was characterised by oligonucleotide fingerprinting (ONFP) to reduce sequencing redundancy. According to the ONFP procedure, clones giving the same hybridisation pattern with a battery of 200 8mer oligonucleotides are grouped into clusters. One

QY 972 CCACTCTTGGCGGTCTGTGCTACTACTGCGGAGCGCGGCGGTGAGCTGCGCA 1031
 DB 942 CGAGAGCTCCGACGTGCACTGCTCTTGGCGGAGCGCGGCGCGCGCGCG 1001
 QY 1032 GGACATTCGACACCGAGCTGAAGCCCGGCGCATGTGCTTCCATGAGCCGATGATGCT 1091
 DB 1002 CAGCTTCACTGAGGACGAGAGGTCTCACTTTGGCAGCCACGACGAGAACCGCTCAT 1061
 QY 1092 GCCGAGGCGCAT-----GCCCGGTGCGCGCGGTATCGCGACGACGACATCTGATGCT 1145
 DB 1062 CGTTCACGACGACATGCGCGCGCGGACCGCGCGCGCGCTCCGACATCCGCGCATGCT 1121
 QY 1146 CGGAGAGACGCTGCGCGGAGACATGACCGGCTTCCGCTGCGGTGCGGAGACAGATCAT 1205
 DB 1122 CAGCTTGTGCGGCGCTGCACTTGTGCGAGCGCGCTACGACGAGTCCACTCTCT 1181
 QY 1206 C 1206
 DB 1182 C 1182

RESULT 11

BZ894814

LOCUS BZ894814 524 bp DNA linear GSS 30-JUL-2003

DEFINITION Hg4_0102 Hg pUC18 Library Halobaculum gomorrense genomic 5', genomic survey sequence.

ACCESSION BZ894814

VERSION BZ894814

KEYWORDS GSS.

SOURCE Halobaculum gomorrense

ORGANISM Halobaculum gomorrense

REFERENCE Archaea; Buryarchaeota; Halobacteriales;

AUTHORS 1 (bases 1 to 524)

Roach, J., Glusman, G., Baliga, N.S., Deutsch, K., Pan, M., Dasgupta, S., Ng, M.V. and Hood, L.

TITLE Low-pass sequencing for microbial comparative genomics

JOURNAL Unpublished (2003)

COMMENT Contact: Gao Y

Institute for Systems Biology

1441 North 34th Street, Seattle, WA 98103, USA

Tel: 206 732 1412

Fax: 206 732 1299

Email: ygo@systemsbiology.org

Seq primer: M13 Forward

Class: shotgun.

FEATURES

source

1..524

Location/Qualifiers

/organism="Halobaculum gomorrense"

/mol_type="genomic DNA"

/strain="ATCC 700876"

/db_xref="taxon:43928"

/clone_lib="Hg pUC18 library"

/note="Vector: pUC18, Site_1: SmaI; A shotgun library was constructed from Halobaculum gomorrense genomic DNA using pUC18/SmaI/BAP plasmid"

ORIGIN

Query Match

Best Local Similarity 49.7%; Score 69.6; DB 8; Length 524;

Matches 177; Conservative 0; Mismatches 179; Indels 0; Gaps 0;

764 TCTTCGCTACTATACACCGCGTGAAGCGACGCTGTTTCGACCATGTGATGACGCCA 823

DB 19 TCGCGGCTACCTCGTGTGTAACAGAGTTCGCGCGGACCTCCAGACGCGGCGGCG 78

QY 824 GCTTGACATCTGGGAGAGAAAGTGGCGGTGATGCGCGGCGGTGCGAGCTGATTAAC 883

DB 79 AGTGGGTTCGCGGAGAGAGCTTCGACGCGGTTCGCGCGGACGCGAGCTGTGTCACG 138

QY 884 CGGCGCGCGGTGCAAGACATGCGCATGAGCTCAACGAGATGTAACGCGAGTGGAGAC 943

DB 139 CCGAGAGTGCAGACACCGGACGACCTCGGATCTACGCGAGGTCAACGCGGAGCCG 198
 QY 944 TGTGAAGTACCGCTCTCTTGGCTATGCGCACTCTTGGGCTGTGCGCACTATGAG 1003
 DB 199 TCGAGAGTCTCTTCAGACGACCAACTCATTTTGGCGTGCAGAGTGTGCTCTTCA 258
 QY 1004 GTCCGAGGCGCGGCGTGGAGCTGCGCGGAGACATGACACCGAGTGAAGCCCGGATG 1063
 DB 259 GCCAGCGCTTACCGCTGAGCGCGGCGGACGCTCATCTTTCACCGGAGACGCGCGGCG 318
 QY 1064 TGTGCTCCAGTACCGGATGATGCTGCTGCGGAGGCGATGCCGCTGCGCGGCG 1119
 DB 319 GGTCTACCGGAGACCGCGGCTGTGCTGCTGAGGCGGACGCGTCACTATCGCG 374

RESULT 12

CB656525

LOCUS CB656525

DEFINITION OSJNEC10P11.f OSJNEC Oryza sativa (japonica cultivar-group) cDNA

ACCESSION CB656525

VERSION CB656525

KEYWORDS EST.

SOURCE Oryza sativa (japonica cultivar-group)

ORGANISM Oryza sativa (japonica cultivar-group)

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Ehrhartoideae; Oryzaceae; Oryza.

1 (bases 1 to 722)

Jantauriyarat, C., Lu, G., Gowda, M., Hatfield, V., Zhou, B., Mazur, E., Kudrna, D., Dean, R., Soderlund, C., Wang, R. and Wang, G.

TITLE Large-scale identification of ESTs involved in the interaction

JOURNAL between rice and Magnaporthe grisea

COMMENT Unpublished (2003)

Contact: Rod Wing

Arizona Genomics Institute

University of Arizona

Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ

85721-0088, USA

Tel: 520 626 3967

Fax: 520 621 9288

Email: http://genome.arizona.edu

PCR Primers

FORWARD: gta aac cga cgg cca gtc

BACKWARD: gga aac agc tat gac cat g

Plate: 10 row: F column: 11

Seq primer: gta aac cga cgg cca gtc.

FEATURES

source

1..722

Location/Qualifiers

/organism="Oryza sativa (japonica cultivar-group)"

/mol_type="mRNA"

/cultivar="Nipponbare"

/db_xref="taxon:39947"

/clone="OSJNEC10P11"

/tissue_type="leaf"

/dev_stage="3 week"

/lab_host="DH10B"

/clone_lib="OSJNEC"

/note="Vector: pBluescript II KS+; Site_1: EcoRI; Site_2: XhoI; 6 hrs after inoculation with Rice Blast (C9240-1)"

ORIGIN

Query Match

Best Local Similarity 45.0%; Score 69.2; DB 6; Length 722;

Matches 260; Conservative 0; Mismatches 318; Indels 0; Gaps 0;

532 GAGCGCGGCGCTGCGCGCTGCGCATGAGCGCGGCTGCGGAGCATGAGTGGCGATC 591

DB 136 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 195

QY 592 GCCACCAACATGCGATGATCGCGAGATGCCCAATGCTTCCCTTGTGAGCTGATG 651

DB 196 AACCTGTCCCGCGCGAGCTGCGCGCGCTCAAGCGCGCGCGCGCGCGCGCGCGCG 255

QY 652 GACACCTGACCTGTTCCAGTGGGATCAACACGACGCGGACAAATCCGCTAC 711
 DB 256 GTGGGCTCGGCGGCGACAGCTGACAGAACCGCCAGGTCTTCTCCCGACCTCC 315
 QY 712 AACCGATCGTGAATCCGCGCATCTCTTTCGCTCAACCTTCGATGATCTTGGC 771
 DB 316 GTGACCTGTTGGGCGAACGCGCTGCTCCGCTCCGACATCAACGCGCTACGGC 375
 QY 772 TACTACACCGGCTGAGCGGACGCTGTTCTGGAACATGTCATGACGCGCTGAC 831
 DB 376 CTGACGCGCTCGACGTCGACCTCAAGCACTTCAACACGACGCGCGCGCTGAC 435
 QY 832 ATCTGGAGAAAGACGTCGCGCTGACATCGCGGCGGCTGAGCTCAAGCGGCGCG 891
 DB 436 ACCTGTGAGAGTCATCGCGCGCTCTCTCAACGAGCTCAAGGCGCGGACCCGAACATC 495
 QY 892 CGCTGACAGACATCGCATCGACGATCAACGAGATGACCGAGTGGGACTGTGTAAG 951
 DB 496 ACCACTCCATCGCGCGCTTCAAGAGACGCGGTGGTGAAGGCTACTACGCGCTGTGG 555
 QY 952 TACCGCTCTTGGGCTATGAGCCACTCTTTCGCGGTGTGCTCACTACGCTGCGAG 1011
 DB 556 CGGCGCTACCGCGCGGTGATGACCTCATCACTTCAAGCTTCAAGGCTACGCGGCAAC 615
 QY 1012 GCCGGGTGAGGCTGGCGGACGATCAACGAGCTGAAGCGCGGCAATGTTGCTTC 1071
 DB 616 ACCGACGTGCGGACGTCATGATGTTCTACGACGAGCGGCGGAATACCGGCGGCG 675
 QY 1072 ATGAGCCGATGATGATGCTGCGCGGAGGCGATGCCGCG 1109
 DB 676 AAGGTCTCACGACGTTCAAGACCGGCGAGCTGCGCGG 713

RESULT 13
 CB683938 722 bp mRNA linear EST 09-APR-2003
 LOCUS OSJUNEf12P21.f OSJUNEf Oryza sativa (japonica cultivar-group) cDNA
 DEFINITION clone OSJUNEf12P21 5', mRNA sequence.

ACCESSION CB683938
 VERSION CB683938
 KEYWORDS EST.
 SOURCE Oryza sativa (japonica cultivar-group)
 ORGANISM Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE 1 (bases 1 to 722)
 Jantaasuriyarat, C., Lu, G., Gowda, M., Hatfield, J., Zhou, B., Mazur, E.,
 Kudrna, D., Dean, R., Soderlund, C., Wing, R. and Wang, G.
 Large-scale identification of ESTs involved in the interaction
 between rice and Magnaporthe grisea
 Unpublished (2003)

JOURNAL COMMENT
 TITLE Arizona Genomics Institute
 CONTACT: Rod Wing
 University of Arizona
 Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
 85721-0088, USA
 Tel: 520 626 3967
 Fax: 520 621 9288
 Email: http://genome.arizona.edu
 PCR Primers
 FORWARD: gta aac cga cgg cca gtc
 BACKWARD: gga aac agc tat gac cat g
 Plate: 12 row: P column: 21
 Seq primer: gta aac cga cgg cca gtc.

FEATURES
 source
 1..722
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="OSJUNEf12P21"

/cissue_type="Leaf"
 /dev_stage="3 week"
 /lab_host="DH10B"
 /clone_id="OSJUNEf"
 /note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
 XhoI; Uninfected Control"

Query Match 5 6%; Score 67.6; DB 6; Length 722;
 Best Local Similarity 44.8%; Pred. No. 0.0002;
 Matches 259; Conservative 0; Mismatches 319; Indels 0; Gaps 0;

QY 532 GGGGCGCGGCGCTGCGCGGCTGCGCATCAAGCGCGGCGTCCCGGACATGATGGGATC 591
 DB 142 GCGGCGGCGCGCACCGCGCGCGCGACCAAGCGGTCTTCTCGCGGATCAAGCGCGGCG 201
 QY 592 GCCACCAACATGCGATGATCGCGGAGATCGCCAAATGCTTCCCTTGTGAGCTGATG 651
 DB 202 AACCTGTCCCGCGCGACGTCGCGCGCGCTCAAGCGCGCGCACCCCAACGCTCAAGCTGATG 261
 QY 652 GACACCTGACCTGTTCCAGTCCGCGCATCAACGAGCGGCGGCGACAAATCCGCTGAC 711
 DB 262 GTGGGCTTGGGCGGCGACGCGTGAAGACACCGCGCAAGGCTTCTTCTCCCGACCTCC 321
 QY 712 AACCGATCGTCAATCCGCGGACATCTTTCGCTCAACCTTCCCGATGATCTTCCG 771
 DB 322 GTGACCTGCTGGGTGGCGCAACCGCGCTCTCGCTCCGCGATCAAGCGCGCTACGGC 381
 QY 772 TACTACACCGGCTGAGGCGACGCTGTTCTGCGACATGTCATGACCGCACTGAC 831
 DB 382 CTGACGCGGCTGACGTCATCAAGCACTTCAACGAGCGCGCGCGCGCTGAC 441
 QY 832 ATCTGGAGAAAGACGTGCGCGCTGATGCGCGCGCGCTGAGCTGATCAAGCGGCGCG 891
 DB 442 ACCTGTGAGAGTCATGCGCGCGCTCTCTACCGAGCTCAAGCGGCGGCACTCCGAACATC 501
 QY 892 CGCTGACAGACATCGCATCGACGCTCAACGAGATGACCGGAGTGGGACTGTGTAAG 951
 DB 502 ACCACTCCATCGCGCGCTTCAAGAGACGCGGTGGTGAAGGCTACTACGCGCTGTGG 551
 QY 952 TACCGCTCTTGGGCTATGAGCCACTCTTTCGCGGTGTGCTCACTACGCTGCGAG 1011
 DB 562 CGGCGCTACGCGCGGTGATGACCTCGTCACTTCAAGCTTCAAGGCTACGCGGCAAC 621
 QY 1012 GCCGGGTGAGGCTGGCGGACGATCAACGAGCTGAAGCGCGGCAATGTTGCTTC 1071
 DB 622 ACCGACGTGCGGACGTCATGATGTTCTACGAGCGGCGGCAATACCGGCGGCGG 681

QY 1072 ATGAGCGGATGATGCTGCGCGGAGGCGATGCCGCG 1109
 DB 682 AAGGTCTCGGACGTTCAAGACCGGCGACATCACCGG 719

RESULT 14
 CL963721 2538 bp DNA linear GSS 21-SEP-2004
 LOCUS OsIFCC038554 Oryza sativa Express Library Oryza sativa (indica
 DEFINITION cultivar-group) genomic, genomic survey sequence.

ACCESSION CL963721
 VERSION CL963721
 KEYWORDS GSS.
 SOURCE Oryza sativa (indica cultivar-group)
 ORGANISM Oryza sativa (indica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE 1 (bases 1 to 2538)
 Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M.,
 Zhao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L.,
 Wong, G. K. S., Deng, X. W. and Wang, J.
 An analysis of transcriptional regulation of the rice genome and
 its comparison to Arabidopsis

JOURNAL
 TITLE Unpublished (2004)

COMMENT Contact: Chen Chen
Department of Bioinformatic
Beijing Institute of Genomics
Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-80488676
Email: chenchen@genomics.org.cn
Rice genomic sequence.
Class: exon-trapped.

FEATURES
source
1..2538
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="genomic DNA"
/db_xref="taxon:39946"
/clone_lib="Oryza sativa Expressed Library"
/note="Oryza sativa exon trapped genomic sequences"

ORIGIN

Query Match 5.4%; Score 66; DB 9; Length 2538;
Best Local Similarity 44.8%; Pred. No. 0.00048; Indels 3; Gaps 1;
Matches 295; Conservative 0; Mismatches 360;

Oy 408 GCTCGAGAGAGCCCTACCGGCGCTCGAGTTCTGTCATCAGCCAGCCCTTCATGTGAT 467
Db 1167 GCGCGTGAAGCGGCTGCTCAGCGCGCTGAGGTGCTCAAGTAGCACTCCGTTGGGAG 1226
Oy 468 GCGCACCATCAAGTGTGCTGAAAGAGCAAGCTGATCCGGAAGCGCGCGTGTGTGA 527
Db 1227 CTTCAAGCGCGAGTTTCGCGTGGACGGAAGAGCGCGACGACGCGAGCGCGCA 1286
Oy 528 CCGTGGCGCGCGCGCTGCGCGCGCTGCCATCAAGCGCGCGCTCCGAGCATGAAGTGC 587
Db 1287 GCGCGTGGGCGGCGGCTGCGAGATGTTTGGCGGCTTCGCGCGCTGGAGCCAT 1346
Oy 588 GATCGCACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 647
Db 1347 GCGCGTGAAGTGTGCTGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1406
Oy 648 GATGAGCATGAGAGCTGTTTCAAGTGGGATCAACCGAGCGCGCGCAATCCGT 707
Db 1407 GCTGCTCCCATCAACCGCGCAATCTTCAACAGAGCGCGGCTCCAGCAAGAG 1466
Oy 708 CACCAACCGCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 767
Db 1467 GCGCTACACCTTCTCTCTCAACCTCGGCTCGGCTCTTCTCTCTCTCTCTCTCTCT 1526
Oy 768 CGGCTACTACCGCGCTGAGCGAGCGAGCTGTTCTGCAACATGTCATGAGCCAGCT 827
Db 1527 GCGCGCGACCAAGAACTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1586
Oy 828 CGACATCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 884
Db 1587 CTACATCGCGAGATGACGAGCGACCAAGATGCGCGTGAAGCGGGAACCGCAGTC 1646
Oy 885 GCGCGCGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 944
Db 1647 GGAGCAAGAGAGATCAAGAGATTCACACCGAGATCGAGTGTCTCAAGCTCCGCCAC 1706
Oy 945 GCTGAAGTACCGCTCTTCTGCTATGAGCACTCTTCTGCGAGTGTGCACTACTACG 1004
Db 1707 CCACTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1766
Oy 1005 TCGCGAGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1062
Db 1767 GTACATGACATAGCGCCCTTCGCGACCACTACGCGAGAGAGAGAGAGAGAGAG 1824

RESULT 15
CL948585 1368 bp DNA linear GSS 21-SEP-2004
LOCUS OIRSB000026 Oryza sativa Expressed Library Oryza sativa (indica
DEFINITION cultivar-group) genomic, genomic survey sequence.
ACCESSION CL948585

VERSION CL948585.1 GI:52360594
GSS
SOURCE Oryza sativa (indica cultivar-group)
ORGANISM Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE
1 (bases 1 to 1368)
Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M.,
Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L.,
Wong, G. K. S., Deng, X. W. and Wang, J.
An analysis of transcriptional regulation of the rice genome and
its comparison to Arabidopsis
Unpublished (2004)

JOURNAL
COMMENT
Contact: Chen Chen
Department of Bioinformatic
Beijing Institute of Genomics
Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-80488676
Email: chenchen@genomics.org.cn
Rice genomic sequence.
Class: exon-trapped.

FEATURES
source
1..1368
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="genomic DNA"
/db_xref="taxon:39946"
/clone_lib="Oryza sativa Expressed Library"
/note="Oryza sativa exon trapped genomic sequences"

ORIGIN

Query Match 5.4%; Score 65; DB 9; Length 1368;
Best Local Similarity 44.4%; Pred. No. 0.00077; Indels 0; Gaps 0;
Matches 263; Conservative 0; Mismatches 330;

Oy 587 CGATGCGCACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 646
Db 137 CGAGAGCGCGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 256
Oy 647 TGATGAGACACTGAGCTGTTTCAAGTGGGATCAACCGAGCGCGCAATCCGT 706
Db 257 TCGGAGACTCACCGCGAGAGCTAGCCCAAGAGCGCGCGCTTGGAGAGAGAGAGAG 316
Oy 707 TCACCAACCGCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 766
Db 317 GTACCAACCGCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 376
Oy 767 TCGGCTACTACCGCGCTGAGCGAGCGAGCTGTTCTGCAACATGTCATGAGCCAGCT 826
Db 377 GCGCGCGTACCGCGCGGAGTGTGCTGCGCGCGCGCGCGCGCTCTGCAAGCGCGCTCG 436
Oy 827 TCGACATCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 886
Db 437 ACATGCGCATCAATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 496
Oy 887 GCGCGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 946
Db 497 TCTGTAGCTCAACAGATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 556
Oy 947 TGAAGTACCGCTCTTCTGCTATGAGCACTCTTCTGCGAGTGTGCACTACTACG 1006
Db 557 TATCTTATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 616
Oy 1007 GCGAGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1066
Db 617 CCAACCGGCTATGAGAGCTGTGCTTCAACCGGTACGCAAAATCAACGCGCCCAAGAACT 676
Oy 1067 TCTTCATGAGAGCGCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCAT 1126
Db 677 TCTTTCGCGCGCTCGCGCGCGCTTCAAGAAATGCGAGAGAGAGAGAGAGAGAGAGAGAGAG 736
Oy 1127 AGCAGCATCTCTGATGTCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1179

Db 737 TGAACGTGCGCGCTGSAAGCGCGCGCTCCGCGACGACGTGTACCAACGCTGTTTC 789

Search completed: July 7, 2005, 23:30:21
Job time : 4430 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: July 7, 2005, 20:27:44 ; Search time 236 Seconds

(without alignments)
8403.252 Million cell updates/sec

Title: US-10-807-228a-2

Perfect score: 1212
Sequence: 1 ATGACTGACGACATGTTGCA.....AACACACATCATCCGCACAC 1212

Scoring table: IDENTITY_NUC

Gapop 10.0 ; Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents NA:*

1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
4: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
5: /cgn2_6/prodata/1/ina/5C_COMB.seq:*
6: /cgn2_6/prodata/1/ina/5D_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1210.4	99.9	1212	1	US-09-940-941-2
2	1210.4	99.9	1212	2	US-08-799-897-2
3	1207.2	99.6	1215	2	US-08-947-726A-1
4	74.8	6.2	858	4	US-09-252-991A-4171
5	74.8	6.2	963	4	US-09-252-991A-4435
6	74.8	6.2	1176	4	US-09-252-991A-4314
7	64	5.6	3157	2	US-08-939-002A-1
8	64	5.3	1294	3	US-09-025-691-2
9	63.8	5.3	18686	4	US-09-902-540-1206
10	62	5.1	2064	1	US-08-343-428-1
11	60.6	5.0	1599	4	US-09-475-515-54
12	60.6	5.0	2112	4	US-09-475-515-56
13	60.6	5.0	2112	4	US-09-475-515-57
14	60.6	5.0	2181	4	US-09-475-515-58
15	60.6	5.0	2634	4	US-09-475-515-64
16	60.6	5.0	4284	4	US-09-902-540-3289
17	60.6	5.0	4766	4	US-09-475-515-73
18	60.6	5.0	17727	4	US-09-902-540-1152
19	60	5.0	1822	4	US-09-949-016-1464
20	60	5.0	1965	4	US-09-220-132-21
21	60	5.0	9399	4	US-09-949-016-13206
22	59.2	4.9	1056	4	US-09-266-965-66
23	59.2	4.9	1293	2	US-08-924-440-1
24	59.2	4.9	53500	4	US-09-266-965-76
25	58	4.8	2712	3	US-09-025-691-4
26	57.8	4.8	3780	4	US-09-902-540-9287
27	57.8	4.8	9873	4	US-09-902-540-9932

28	57.4	4.7	1434	4	US-09-902-540-4622	Sequence 4622, Ap
29	57.4	4.7	6776	4	US-09-902-540-884	Sequence 884, Ap
30	57.4	4.7	14809	4	US-09-902-540-1032	Sequence 1032, Ap
31	57.2	4.7	1695	4	US-09-902-540-4319	Sequence 4319, Ap
32	57.2	4.7	2277	4	US-09-902-540-9434	Sequence 9434, Ap
33	57.2	4.7	11358	4	US-09-902-540-1075	Sequence 1075, Ap
34	57.2	4.7	12730	4	US-09-902-540-1029	Sequence 1029, Ap
35	56.6	4.7	1476	3	US-09-434-288-12	Sequence 12, Ap
36	56.4	4.7	876	4	US-09-902-540-5288	Sequence 5288, Ap
37	56.4	4.7	1005	4	US-09-252-991A-1549	Sequence 1549, Ap
38	56.4	4.7	1590	4	US-09-252-991A-1617	Sequence 1617, Ap
39	56.4	4.7	1659	4	US-09-252-991A-1566	Sequence 1566, Ap
40	56.4	4.7	30783	4	US-09-902-540-1258	Sequence 1258, Ap
41	55.8	4.6	1644	4	US-09-902-540-9071	Sequence 9071, Ap
42	55.8	4.6	7719	4	US-09-902-540-969	Sequence 969, Ap
43	55.4	4.6	651	4	US-09-902-540-4369	Sequence 4369, Ap
44	55.4	4.6	2655	1	US-08-471-033-17	Sequence 17, Ap
45	55.4	4.6	2655	1	US-08-471-033-26	Sequence 26, Ap

ALIGNMENTS

RESULT 1
US-09-940-941-2
Sequence 2, Application US/09940941
Patent No. RE38687
GENERAL INFORMATION:
APPLICANT: Sogabe, Atsushi
Hattori, Takashi
Nishiyama, Yoshiaki
Kawamura, Yoshiaki
TITLE OF INVENTION: NOVEL CREATINE AMIDINOHYDROLASE, PRODUCTION
THEREOF AND USE THEREOF
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: LEYDIG, VOIT & MAYER, LTD.
STREET: Two Prudential Plaza, Suite 4900
CITY: Chicago
STATE: Illinois
COUNTRY: US
ZIP: 60601-6780
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/940,941
FILING DATE: 28-Aug-2001
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/799,897
FILING DATE: 13-FEB-1997
APPLICATION NUMBER: JP 25435/1996
FILING DATE: 13-FEB-1996
ATTORNEY/AGENT INFORMATION:
NAME: Robert F. Green
REGISTRATION NUMBER: 27555
REFERENCE/DOCKET NUMBER: 78064
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 616-5600
TELEFAX: (312) 616-5700
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1212 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA
ORIGINAL SOURCE:
ORGANISM: Alcaligenes faecalis

STRAIN: TE3581 (FERM P-14237)
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1 to 1212
 SEQUENCE DESCRIPTION: SEQ ID NO: 2:
 US-09-940-941-2

Query Match 99.9%; Score 1210.4; DB 1; Length 1212;
 Best Local Similarity 99.9%; Pred. No. 1.8e-242;
 Matches 121; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 ATGACTGACGACATGTTGACGTTGATGAAATGACCAACGCGGAGAAAGATTATTCGCG 60
DB 1 ATGACTGACGACATGTTGACGTTGATGAAATGACCAACGCGGAGAAAGATTATTCGCG 60
QY 61 TTTTCGGATGCGGAGATGACCGCGCGCAAAAGAGTTGCGGCTGATGGCCAAAG 120
DB 61 TTTTCGGATGCGGAGATGACCGCGCGCAAAAGAGTTGCGGCTGATGGCCAAAG 120
QY 121 AATGTCGATGCGGCGCTGTTCACTCTTATCACTGATCACTACTATTCGCGCTGCTG 180
DB 121 AATGTCGATGCGGCGCTGTTCACTCTTATCACTGATCACTACTATTCGCGCTGCTG 180
QY 181 TACTGCTATTTGCGACGCAAGTACGCGCATGTCATGACCAACAGCGCAAGAGAT 240
DB 181 TACTGCTATTTGCGACGCAAGTACGCGCATGTCATGACCAACAGCGCAAGAGAT 240
QY 241 TCGGCGCGGATCGACGCGCGCGACGCTTGCGCGCGGAGCTTCGCGCAACATCACTAC 300
DB 241 TCGGCGCGGATCGACGCGCGCGACGCTTGCGCGCGGAGCTTCGCGCAACATCACTAC 300
QY 301 ACCGACTGCGCGCGCGCAATTTCTATCGCGCGCTGCGGCACTGACCAACGCGCGCG 360
DB 301 ACCGACTGCGCGCGCGCAATTTCTATCGCGCGCTGCGGCACTGACCAACGCGCGCG 360
QY 361 CGCATCGGATGAGTTGACCAACGCACTTCGCGCGCGAGCTCGAGGAAAGCC 420
DB 361 CGCATCGGATGAGTTGACCAACGCACTTCGCGCGCGAGCTCGAGGAAAGCC 420
QY 421 CTACCGGCGGTCGATGCTTCGACATCAACGCGCGCTGATGATGCGCAACATCAG 480
DB 421 CTACCGGCGGTCGATGCTTCGACATCAACGCGCGCTGATGATGCGCAACATCAG 480
QY 481 TCGCTCGAAGACAGAGCTGATTCGCGGAGGCGCGCTGTGTGACGTGCGCGCGCG 540
DB 481 TCGCTCGAAGACAGAGCTGATTCGCGGAGGCGCGCTGTGTGACGTGCGCGCGCG 540
QY 541 GCGTCGCGGCGGTCGACATCAAGGCGCGCGCGCGGATGAGATGCGCATGCGCAAC 600
DB 541 GCGTCGCGGCGGTCGACATCAAGGCGCGCGCGCGGATGAGATGCGCATGCGCAAC 600
QY 601 AATGCGATGATCCGCGAGATCCGCAATGCTTCCCTTCGTGAGCTGATGACACTGCG 660
DB 601 AATGCGATGATCCGCGAGATCCGCAATGCTTCCCTTCGTGAGCTGATGACACTGCG 660
QY 661 ACCTGCTTCAGTCCGCGCATCAACACGCGCGCGCGCAATTCGCGTCAACACGCGATC 720
DB 661 ACCTGCTTCAGTCCGCGCATCAACACGCGCGCGCGCAATTCGCGTCAACACGCGATC 720
QY 721 GTCGCAATCCGCGGACATCTTTCGTCGACACCTTCGCGATGATCTTCGCGCTACTAC 780
DB 721 GTCGCAATCCGCGGACATCTTTCGTCGACACCTTCGCGATGATCTTCGCGCTACTAC 780
QY 781 GCGCTGAGACGACGCTGTTCTGCGACATGTCGATGACGCGAGCTTCGACATCTGGAG 840
DB 781 GCGCTGAGACGACGCTGTTCTGCGACATGTCGATGACGCGAGCTTCGACATCTGGAG 840
QY 841 AAGAACTGCGCGTGCATGCGCGCGGCTGACGCTGATCAAGCGCGCGCGCGCTGCAAG 900
DB 841 AAGAACTGCGCGTGCATGCGCGCGGCTGACGCTGATCAAGCGCGCGCGCGCTGCAAG 900
QY 901 GACATCCGATCGAGCTCAACGAGATGATCCGCGAGTGGAGCTGCTGAAGTACGCTCC 960
  
```

```

DB 901 GACATCCGATCGAGCTCAACGAGATGATACCGGAGTGGACCTGCTGAAGTACCGCTCC 960
QY 961 TTGCGCTATGGCCACTCTTTCGCGGCTGTGTGCCACTATCTACGCTGCGAGCGCGCTG 1020
DB 961 TTGCGCTATGGCCACTCTTTCGCGGCTGTGTGCCACTATCTACGCTGCGAGCGCGCTG 1020
QY 1021 GAGCTGCGGAGACATCGACACCGAGCTGAAGCGCGGATGATGCTTCATGAGCGCG 1080
DB 1021 GAGCTGCGGAGACATCGACACCGAGCTGAAGCGCGGATGATGCTTCATGAGCGCG 1080
QY 1081 ATGTCGATGCTGCGGAGGCGATGCGCGGCGGCTATTCGACGACGACATCTCTG 1140
DB 1081 ATGTCGATGCTGCGGAGGCGATGCGCGGCGGCTATTCGACGACGACATCTCTG 1140
QY 1141 ATTCGTCGGGAGGACGCTGTCGAGAACATCAACGCGCTTCGCTGCTGCGGAAACAAC 1200
DB 1141 ATTCGTCGGGAGGACGCTGTCGAGAACATCAACGCGCTTCGCTGCTGCGGAAACAAC 1200
QY 1201 ATCATCCGCAAC 1212
DB 1201 ATCATCCGCAAC 1212
  
```

RESULT 2
 US-08-799-897-2
 Sequence 2, Application US/08799897
 Patent No. 6080553
 GENERAL INFORMATION:
 APPLICANT: Sogabe, Atsushi
 APPLICANT: Hattori, Takashi
 APPLICANT: Nishiyama, Yoshiaki
 APPLICANT: Kawamura, Yoshiaki
 TITLE OF INVENTION: NOVEL CREATINE AMIDINOHYDROLASE, PRODUCTION
 NUMBER OF SEQUENCES: 3
 TITLE OF INVENTION: THERBOF AND USE THEREOF
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: LEYDIG, VOIT & MAYER, LTD.
 STREET: Two Prudential Plaza, Suite 4900
 CITY: Chicago
 STATE: Illinois
 COUNTRY: US
 ZIP: 60601-6780
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/799,897
 FILING DATE: 13-FEB-1997
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 25435/1996
 FILING DATE: 13-FEB-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Robert F. Green
 REGISTRATION NUMBER: 27555
 REFERENCE/DOCKET NUMBER: 78064
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (312) 616-5600
 TELEFAX: (312) 616-5700
 TELEX: 25-3533
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1212 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: genomic DNA
 ORIGINAL SOURCE:
 ORGANISM: Alcaligenes faecalis
 STRAIN: TE3581 (FERM P-14237)
 FEATURE:

US-08-947-726A-1

Query Match: 99.6%; Score 1207.2; DB 2; Length 1215;
Best Local Similarity 99.8%; Pred. No. 8.4e-242;
Matches 1209; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
QY 1 ATGACTGACGACATGTTGACAGTGAATAAGGCAACGCGGAGAAAGATTATTCGCGG 60
DB 1 ATGACTGACGACATGTTGACAGTGAATAAGGCAACGCGGAGAAAGATTATTCGCGG 60
QY 61 TTTTCGGATGCGGAGATGACCGCGCCGCAAAAGAGCGTTGCGGCGTGAATGGCAAGAC 120
DB 61 TTTTCGGATGCGGAGATGACCGCGCCGCAAAAGAGCGTTGCGGCGTGAATGGCAAGAC 120
QY 121 AATGTGATGCGGCGCTGTTCACTTCTATCACTGATCACTAATTCGCGTGGCTG 180
DB 121 AATGTGATGCGGCGCTGTTCACTTCTATCACTGATCACTAATTCGCGTGGCTG 180
QY 181 TACTGTATTTGCGAGCGCAAGTACGCGCATGATGACCAACAAGCGCAGACGATT 240
DB 181 TACTGTATTTGCGAGCGCAAGTACGCGCATGATGACCAACAAGCGCAGACGATT 240
QY 241 TCGGCGCGGATGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
DB 241 TCGGCGCGGATGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
QY 301 ACCGACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
DB 301 ACCGACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
QY 361 CGCATCGGATGAGATTCACCAACGCGCATTCGCGCGCGCGCGCGCGCGCGCGCGCG 420
DB 361 CGCATCGGATGAGATTCACCAACGCGCATTCGCGCGCGCGCGCGCGCGCGCGCGCG 420
QY 421 CTACCGCGCGTGAATTCGTCGACATCAAGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
DB 421 CTACCGCGCGTGAATTCGTCGACATCAAGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
QY 481 TCGCTGGAAGAGAGAGTGTGCGCGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
DB 481 TCGCTGGAAGAGAGAGTGTGCGCGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
QY 541 GCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
DB 541 GCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
QY 601 AATGCGATGATCGCGAGATCGCGCAATCGTTCCCTTCGTGAGCTGATGACACCTGG 660
DB 601 AATGCGATGATCGCGAGATCGCGCAATCGTTCCCTTCGTGAGCTGATGACACCTGG 660
QY 661 ACTGCGTTCAGTGGGCGATCAACACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
DB 661 ACTGCGTTCAGTGGGCGATCAACACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
QY 721 GTGCGATTCGCGCGACATCCCTTCGTCACACCTTCGCGATGATCTTCGCGTACTAC 780
DB 721 GTGCGATTCGCGCGACATCCCTTCGTCACACCTTCGCGATGATCTTCGCGTACTAC 780
QY 781 GCGCTGAGCGCGCGCTGTTCTGCGACATGTCGATGACGCGCGCGCGCGCGCGCGCG 840
DB 781 GCGCTGAGCGCGCGCTGTTCTGCGACATGTCGATGACGCGCGCGCGCGCGCGCGCG 840
QY 841 AAGAAAGTGGCGCTGATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
DB 841 AAGAAAGTGGCGCTGATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
QY 901 GACATGCGCATGAGCTCAACGAGATGTAACGCGAGTGGGAGCTGCTGAAGTACCGCTCC 960
DB 901 GACATGCGCATGAGCTCAACGAGATGTAACGCGAGTGGGAGCTGCTGAAGTACCGCTCC 960
QY 961 TTTCGCGTATGCGCATCTCTTCGCGCGCTGCTGTGACACTACGCTGGCGAGCGCGCG 1020
DB 961 TTTCGCGTATGCGCATCTCTTCGCGCGCTGCTGTGACACTACGCTGGCGAGCGCGCG 1020
```

```
QY 1021 GAGCTGCGGAGAGATGACATGACACCGAGTGAAGCCCGGCGATGATGTCATGAGCGG 1080
DB 1021 GAGCTGCGGAGAGATGACATGACACCGAGTGAAGCCCGGCGATGATGTCATGAGCGG 1080
QY 1081 ATGTGATGATGCTCCGAGAGGAGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1140
DB 1081 ATGTGATGATGCTCCGAGAGGAGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1140
QY 1141 ATGTGCGGAGAGAGCGTGGCGGAGACATCAACCGCGCTTCGCTTGGCTCCGGAACAC 1200
DB 1141 ATGTGCGGAGAGAGCGTGGCGGAGACATCAACCGCGCTTCGCTTGGCTCCGGAACAC 1200
QY 1201 ATCATCCGCGAC 1212
DB 1201 ATCATCCGCGAC 1212
```

RESULT 4

```
US-09-252-991A-4171
; Sequence 4171, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.116
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 4171
; LENGTH: 858
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-4171
```

Query Match: 6.2%; Score 74.8; DB 4; Length 858;
Best Local Similarity 45.7%; Pred. No. 1.4e-06;
Matches 339; Conservative 0; Mismatches 397; Indels 6; Gaps 2;

```
QY 471 CACCATCAAGTGTCTCGAAGAGCAGAGCTGATCCGGAAGGCGCGCGCGCGCGCGT 530
DB 81 CACCATCAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 140
QY 531 CGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 590
DB 141 AGTGTGGAATGATTCGCGGAACACATCAAGCGCGCGCGCGCGCGCGCGCGCGCG 200
QY 591 CGCGACCAACCAATGCGATGATCGCGAGATCGCGCAATGCTTCCTTCGTGAGCTGAT 650
DB 201 CATTCGCCACGACTATATGTCATCAAGAGAGCGAGATTCGCGCGCGCGCGCGCGCG 260
QY 651 GGAACCTGGAACCTGTTCCAGTCCGCGATCAACACCGAGCGCGCGCGCGCGCGCG 710
DB 261 GGGCTTTCCCAAGTGAATTCGATCTCGATCAACCATGATGATGCTGCAATGCGCTCA 320
QY 711 CAACGCGATTCGCGATTCGCGCGAGATCTTCCTTCGCTCAACATTCCTCCGATATCTT 770
DB 321 CGAAGAGCGCTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 380
QY 771 CTACTACACCGCGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 830
DB 381 CTACACAGCGCGCACACAGCAAGATGTTCTGCTGCGCAAGACCGCGAGATGAGCGCG 440
QY 831 CATTCGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 890
DB 441 CCTTCGCGAGATCACCGAGAGATGATCAAGAGCAATTCCTGCTGCTGCTGCTGCTG 500
QY 891 GCGCTGCAAGAGATTCGCGATGAGCTCAACGAGATGTAACCGCGAGTGGAGCTGCGAA 950
```

```

Db      501 GCACCTGCGCGGATATCG---GCGAAATCATCCAGAAAGCAGCGAAAGAAAGCGCTTCTC 557
Qy      951 GATACCGCTCTTGGGTATATGACCATCTCTTGGCGGTCTGTGCACTATACGTGTGCCA 1010
Db      558 GGTGTCTCCGGAATCTATGCGCGCATGCGCATGCGCAAGGTGTTCC---ACGAGAGCGCGA 614
Qy      1011 GCGCGCGGTGAGCTGCGCGAGGACATCGACCGAGCTGAAGCCCGGCAATGTTGTCTC 1070
Db      615 GGTCTTCCACTACGTCTGTGTGCGGATACCGGATCTGAGGAAAGGATATCTTCAAC 674
Qy      1071 CATGAGCCGATGTGTATGTCTGCGGAGGCGATGCCGTGCGCGGTATCTGAGAGA 1130
Db      675 CATCGAGCCGATGTATCAACAGAGGCGCGCGGAGACCGCGCTGCTGCGGAGCGGTGAGC 734
Qy      1131 CGACATCTCTGATGTGTGTGCGGAGGACGCGTCCCGAATCATACCGGCTTCCCTTGGTGC 1190
Db      735 CGCGATCACCAAGAGCCGCAAGCTGTCCGCGCATGTAGAGACACCGTGTCTGTACCGC 794
Qy      1191 GGAACACAACATCATCCGCAAC 1212
Db      795 CGATGGCTTACGAGATCTCTAC 816

```

RESULT 5

```

US-09-252-991A-4435/c
; Sequence 4435, Application US/09252991A
; Patent No. 6551795

```

GENERAL INFORMATION:

```

; APPLICANT: Marc J. Rubenfield et al.

```

```

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136

```

```

; CURRENT APPLICATION NUMBER: US/09/252,991A

```

```

; CURRENT FILING DATE: 1999-02-18

```

```

; PRIOR APPLICATION NUMBER: US 60/074,788

```

```

; PRIOR FILING DATE: 1998-02-18

```

```

; PRIOR APPLICATION NUMBER: US 60/094,190

```

```

; PRIOR FILING DATE: 1998-07-27

```

```

; NUMBER OF SEQ ID NOS: 33142

```

```

; SEQ ID NO 4435

```

```

; LENGTH: 963

```

```

; TYPE: DNA

```

```

; ORGANISM: Pseudomonas aeruginosa

```

```

US-09-252-991A-4435

```

```

Query Match      6.2%; Score 74.8; DB 4; Length 963;
Best Local Similarity 45.7%; Pred. No. 1.4e-06;
Matches 339; Conservative 0; Mismatches 397; Indels 6; Gaps 2;

```

```

Qy      471 CACCATCAAGTGTCTGGAAGACAGAACTGATCCGCGAAGCGCCGCTGTGTGACT 530
Db      922 CACCATCAAGAGCGCCGAGATATCGAAGATGCGCATGCGCGCGCTGCGCGCGA 863
Qy      531 CGCGGCGCGCGCTGCGCGCTGCGATCAAGCGCGCGCGCGCGAGCATGAATGCGCAT 590
Db      862 AGTGTGGAATATGATGCGGAAACATCAAGCGCGCGCGCGCGAGCATGAATGCGCAT 803
Qy      591 CGCGCAACCAATGAGATATCGCGAGATCGCCCAATGCTTCCCTTGTGAGCTGAT 650
Db      802 CATCTGCGCAAGATATGATGATCAAGAGAGAGGAGATCCCGCCCTTGAATCAACA 743
Qy      651 GGAACACTGAGACTGTGCTCACTGCGGATCAACCGAGCGCGCGCGCAATCCGATCAC 710
Db      742 GGGCTTTCCTCAAGTGTGATGCACTCGATCAACATGTGTCTGCGATGCGATCCCA 683
Qy      711 CAACCGCATGTGAGATCGGAGATCTCTTGTGCTCAACCTTCCGATGATCTTGG 770
Db      682 CGAGAGCGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 623
Qy      771 CTACTACACCGCGCTGAGAGCGAGCGCTGTCTGCGACCATGTGATGAGCGCGAGCTCGA 830
Db      622 CTACAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 563

```

```

Qy      831 CATCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 890
Db      562 CTTCTGCGAGATACCGAGAGATGATGATCAAGAGGATTTCCGTGTGTCTGCGCGCG 503
Qy      891 GCGCTGCAAGAGATGCGCATGAGCTCAAGAGATGATGATGATGATGATGATGATGAT 950
Db      502 GCACTGCGAGAGATGCG---GCGAATATCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 446
Qy      951 GATACCGCTCTTGGGTATATGACCATCTCTTGGCGGTCTGTGTGCACTATACGTGTGCCA 1010
Db      445 GGTGTCTCCGGAATCTGTGCGCGCATGCGCATGCGCAAGGTGTTCC---ACGAGAGCGCGA 389
Qy      1011 GCGCGCGGTGAGCTGCGCGAGAGATCGACACCGAGCTGAAGCCCGGCAATGTTGTCTC 1070
Db      388 GGTCTTCCACTACGTGTGTGCGGATACCGGATCTGAGAGAGAGAGAGATATCTTCAAC 329
Qy      1071 CATGAGCCGATGTGTATGTCTGCGGAGGCGATGCCGTGCGCGGTATCTGAGAGA 1130
Db      328 CATCGAGCGGATGATCAACAGAGGCGCGCGGAGACCGCGCTGCTGCGGAGCGGTGAGC 269
Qy      1131 CGACATCTCTGATGTGTGTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1190
Db      268 CGCGATCACCAAGAGCCGCAAGCTGTCCGCGCATGTAGAGACACCGTGTGTGTACCGC 209
Qy      1191 GGAACACAACATCATCCGCAAC 1212
Db      208 CGATGGCTTACGAGATCTCTAC 187

```

RESULT 6

```

US-09-252-991A-4314
; Sequence 4314, Application US/09252991A
; Patent No. 6551795

```

GENERAL INFORMATION:

```

; APPLICANT: Marc J. Rubenfield et al.

```

```

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136

```

```

; CURRENT APPLICATION NUMBER: US/09/252,991A

```

```

; CURRENT FILING DATE: 1999-02-18

```

```

; PRIOR APPLICATION NUMBER: US 60/074,788

```

```

; PRIOR FILING DATE: 1998-02-18

```

```

; PRIOR APPLICATION NUMBER: US 60/094,190

```

```

; PRIOR FILING DATE: 1998-07-27

```

```

; NUMBER OF SEQ ID NOS: 33142

```

```

; SEQ ID NO 4314

```

```

; LENGTH: 1176

```

```

; TYPE: DNA

```

```

; ORGANISM: Pseudomonas aeruginosa

```

```

Query Match      6.2%; Score 74.8; DB 4; Length 1176;
Best Local Similarity 45.7%; Pred. No. 1.5e-06;
Matches 339; Conservative 0; Mismatches 397; Indels 6; Gaps 2;

```

```

Qy      471 CACCATCAAGTGTCTGGAAGACAGAACTGATCCGCGAAGCGCCGCTGTGTGACT 530
Db      244 CACCATCAAGAGCGCCGAGATATCGAAGATGCGCATGCGCGCGCTGCGCGCGA 303
Qy      531 CGCGGCGCGCGCTGCGCGCTGCGATCAAGCGCGCGCGCGCGAGCATGAATGCGCAT 590
Db      304 AGTGTGGAATATGATGCGGAAACATCAAGCGCGCGCGCGCGAGCATGAATGCGCAT 363
Qy      591 CGCGCAACCAATGAGATATCGCGAGATCGCCCAATGCTTCCCTTGTGAGCTGAT 650
Db      364 CATCTGCGCAAGATATGATGATCAAGAGAGAGGAGATCCCGCCCTTGAATCAACA 423
Qy      651 GGAACACTGAGACTGTGCTCACTGCGGATCAACCGAGCGCGCGCAATCCGATCAC 710
Db      424 GGGCTTTCCTCAAGTGTGATGCACTCGATCAACCATGTGTCTGCGATGCGATCCCA 483
Qy      711 CAACCGCATGTGAGATCGGAGATCTCTTGTGCTCAACCTTCCGATGATCTTGG 770

```

```

Db      484 CGAGAGCCGCTGAGAGAGGAGACATCTTCAACGTCCGACATCACCGTCATCAAGACCG 543
      771 CTACTACACCGCGCTGAGAGCGACCGCTGTCTTCCGACCAATGTGATGACCGACGCTCGA 830
      544 CTACACCGCGGACACCGAGCAAGATGTCTTGTGCGGACAGACCGGAAATGCGCGACCG 603
      831 CATCTGGAGAGAAACGTGCGCGTGCATGCGCGCGGCTCGAGCTGATCAACCGCGCGC 890
      604 CTTGCGCCGATCACCAGGAATGATGATCAAGAGGCAATTTCCGTGCGCGCTCGCGCGC 663
      891 GCGCTGCAAGACATCGCCATGAGCTCAACGAGATGACCGCGAGTGGACCTGCTGNA 950
      664 GCACTGCGCGCATATCG--GCGAATATCTCAAGACACGCGGAAAGAGAGCGCTTCC 720
      951 GTACCGCTCTTCCGCTATGCGCACTCTTCCGCGCTGTGCTGCTCACTACATGCTCGCA 1010
      721 GGTGCTCGCGAATATCTGCGCGCATGCGCATCGGCAAGGTGTTCC--ACGAGAGCGCGCA 777
      1011 GCGCGCGCTGAGCTGCGCGGAGACATGACACCGAGCTGAAGCCCGCATGTGTCTC 1070
      778 GGTCTCTCACTACGCTGTGCGCGGATACCGGATGAGAGAGAGGATATCTTCC 837
      1071 CATGAGCCGATGATGATGATCTGCGGAGGAGCATGCGCGGCTGCGGCTATCGCGAGA 1130
      838 CATGAGCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 897
      1131 CGACATCTCTGATGCTGCGGAGAGACGCTGCGAGAAATCATCACCGCTTCCGCTGCTC 1190
      898 CGCGATTCACCAAGACCGCAAGCTGTCCGCGCAATGAGAGACACCGTGTGTCACCGC 957
      1191 GGAACACAAATCATCTCCGCAAC 1212
      958 CGATGCTACGAGATCTTCCACC 979

```

```

RESULT 7
US-08-939-002A-1
; Sequence 1, Application US/08939002A
; Patent No. 5849529
; GENERAL INFORMATION:
; APPLICANT: HAYASHI, KIYOSHI
; APPLICANT: LIU, AIMIN
; APPLICANT: LI, HEBIAO
; APPLICANT: HARAGUCHI, KAZUTOMO
; APPLICANT: KITAMURA, YOSHIAKI
; TITLE OF INVENTION: CELLULOSE PHOSPHORYLASE GENE, VECTOR
; TITLE OF INVENTION: AND TRANSFORMANT CONTAINING SAID GENE
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
; CITY: ARLINGTON
; STATE: VA USA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/939,002A
; FILING DATE: 26-SEP-1997
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: JP 221193/1997
; FILING DATE: 04-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 8361-001-0

```

```

TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3157 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Cellvibrio gilvus
; STRAIN: ATCC 13127
; INDIVIDUAL ISOLATE: Direct Origin: pUC-2
;
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 359..2824
; OTHER INFORMATION: /note="METHOD FOR DETERMINING
; OTHER INFORMATION: SEQUENCE: E"
;
; NAME/KEY: CDS
; LOCATION: 359..2824
;
; US-08-939-002A-1
;
; Query Match 5.6%; Score 68; DB 2; Length 3157;
; Best Local Similarity 42.3%; Pred. No. 4.6e-05;
; Matches 438; Conservative 0; Mismatches 595; Indels 3; Gaps 1;
;
      180 GTACTGCTATTTCGAGCGCAAGTACGGCATGTGATGACCAACAACAGCGACGACAT 239
      802 GTTCGTGCAAGTTCTGCTGTGGAACGCGACAGACGACCAAGCACTACAGCGCACTT 861
      240 TTCGCGCGGATCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 299
      862 GTGATGCGCGAGATCGAGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 921
      300 CACGCGATGCGCGCGCGCGCGCAATTTCTATCGCGCGCGCGCGCGCGCGCGCGCGCGCG 359
      922 GACCGAGTACCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 981
      360 GCGCATGCGCATCGAGTTGACCAACGTCATCTGATCTTCCGCGCGCGCGCGCGCGCGCG 419
      982 CGGCTTGCACACGAGACCGCGCACAGTCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1041
      420 CCTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 479
      1042 GTCGCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1101
      480 GTGCTCGAAGAGAGAGATGATCCGCGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 539
      1102 GCACTCCGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1161
      540 GGCCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 599
      1162 CTACCTGGAAGAACCCGACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1221
      600 CAATCGCATGATCCGCGGATCGCGCAATCGTCCCTTGGAGAGTGAATGACACCTG 659
      1222 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1281
      660 GACTGTTTCAAGTTCGCGCATCAACACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 719
      1282 GAGAGCGGTGAAGTCTACTGAGAGCACTGTCTCGAGTACTGAGTGTGAGACCGGA 1341
      720 CGTGAATCGCGCGCATCTTTCGCTCAACACTTCCGATGATCTTCCGTTACTACAC 779
      1342 CGAGAGGCTCGACCGGATGTCACATCTGGAACCAAGTACAGTACGATGTCACGTTCA 1401
      780 CGCGTGAAGGACACGCTGTTCTGAGACCATGTGATGACGCGACCTGACATCTGGA 839
      1402 CATGTGCGCTGCGGCGTGTCTTCTGAGACGAGGATCGCGCGCGGATGCGCTTCCGGA 1461
      840 GAAGAACGTGCGCGGTGATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 896

```


QY	242	CGGCGCGGCATCGACGGGCGGCACGCTGGGGCGCGAGCTTTCGGCGACAAACATCACTTCA	301
Db	13107	CGCGCGCGCGCTTTCGGGCGCCCAAGACACTGCTGTGTGTGTGACAGTGGCACAAAGTACA	13048
QY	302	CCGACTGCGCGCGCGACAAATTTCTATCGCGCGCTGCGCGACGCTGACACAGCGCGCGCAAGC	361
Db	13047	AGAAAGTCCACCGCGCGCGGATGCGCGCGGAGCGCGCAAGAGGGCCCGAGAGTGTTC	12988
QY	362	GCAATCGGCATCGAGTTTCGACCAACGTCATCTCGACTTTCGCGCGCACGCTCGAGGAAGCC	421
Db	12987	GCAAGGGCATTCGCGCGCGGGGTATCAGCGCGCGCGGAGGTGCGCTCCACATCCCC	12928
QY	422	TACCGGCGCTGCGAGTTCTGTGACATCAAGCAGCCCTTCGATGTGTGATGCGACCATCAAGT	481
Db	12927	GTCCTGGACTTCGCGCTGACCGGCGCGCGCTCGCGCAAGACCGCGGCTTCGACATCAAGA	12868
QY	482	CGCTCGAAGACGAAGAGCTGATCCGGGAAGCGCGCGGTGTGTGACGTGCGAGCGCGCG	541
Db	12867	CGCGGACGTCATTCGCGCGCATTCGCGACGAGCGCTGCAAGCGCGCGCGGAGTGTGTGCAAG	12808
QY	542	CTGCGCGGCTGTCGATTCAGAGCGGCGGTGCGCGAGCATGAATGTGCGATCGCACACCA	601
Db	12807	AGGTGCGCTCCACAGTCCGCTCCGGGCATCAACAGGACGAGCTGTGATGGAATTAACACG	12748
QY	602	ATGCGATGATCCGCGAGATCGCAAAATCGTTCCCTTGTGTGAGCTGATGACACTGGA	661
Db	12747	AGGGGTACATCAAGCGGGGGGTACCCAGCACGCTCAATACACCGCTACCGGAAGT	12688
QY	662	CTGTGTTCCAGTGGGGCATCAACCGAGCGCGCGCAATTCGGTCAACCAACCGCATCG	721
Db	12687	CGCTTCGCAAGTCCGTCACAGAGTCACTGTCACGCGCATCCCGGA---CACCGGCGCG	12631
QY	722	TGCAATCCGGCGCATCTTCTGCTCAACCTTCCGATGATCTTCCGCTACTACAGG	781
Db	12630	TGGAGACCGCGATATGTCMACTTGACCTTCACTCACTTACTGTGTGTGACAGGTG	12571
QY	782	CGCTGAGCGCACGCTGTTCTGCGACCATTCGATGACCGCACGCTCGACATCTGGAGAA	841
Db	12570	ACTGCTCGGCGACTTATCTGTGTGGCAACTGTGATGTAGAGAGCCAGCGGCTGTGTGCGG	12511
QY	842	AGAACTGTGCGCTGTCATCGCGCGGCGCTGAGCTGATCAAGCGGCGCGCGCTGCAAG	901
Db	12510	TGACGCGCGAGTGTCTGTAGACAGGCAATCGCGCGGTGAAGCGGCGCGGCATACAGG	12451
QY	902	ACATCGGCATCGAGCTCAACGAATGTACCGCGAGTGTGACCTGTCTGAAGTACCGTCTCT	961
Db	12450	ACATTCGCGCGCGCATTTGAACGATTCACCCAGACCGGCATGAGCGTGTGTGCGGCT	12391
QY	962	TTCGCTATGCGCACTCTTTCGCGGTGTGTGTGCGCACTTACCGGTGCGAGCGCGGCTGTG	1021
Db	12390	ACTTCC---GGCCACGCGCATCGGCGAAGCTTTCACACACTTGTCTTCAGTGTCCGCACTACT	12334
QY	1022	AGCTGCGCGAGGACATTCGACACGAGCTGGAAGCCGCGCATGTGTGTCTTCATGTAGACCGA	1081

LOCATION: 435
IDENTIFICATION

US-08-343-428-1

Query Match 5.1%; Score 62; DB 1; Length 2064;
 Best Local Similarity 44.4%; Pred. No. 0.00075;
 Matches 342; Conservative 0; Mismatches 420; Indels 8; Gaps 2;

```

QY 103 GCGTGGATGCGCAAGAACATATGTCATGTCGCGCTTTCATCTTATCATGATCAAC 162
DB 993 GCTTCAAGTCAACCAAGAACGCGCTTCTCTATCTTCAACCGCGGCACTGCAACAC 1052
QY 163 TACTATTCGCGCTGCTGTAATCTATTTTGGACGCAAGTACGGCATGTCATCGAAC 222
DB 1053 CTCTGTCACCTGCTGCTCAAC---TCGGGGGACCTTCATCGGCTCCGGGAGGAC 1109
QY 223 AACAAAGCAACGCAATTTCCGCGCATGACGGGCGCAAGCCCTGGGCGCGCACTTC 282
DB 1110 ACCACCTTCCCAACCAACGCACTACGGCATGTCCTGCTACCAACCAACCAACGCA 1169
QY 283 GCGCAACATCACTCAACCGACTGGCGCGCAATTTTCTATCGCGCGGCGGACG 342
DB 1170 GCGCGGGTCACTGTACACGCGCGCTACAGACATGCTCCGCGGCGCAAGCGCTTC 1229
QY 343 CTGACCAAGCGGCGCAAGGCAATCGGCATCGAGTTCAGCAACATCTGACTTCCGC 402
DB 1230 GTGGGCGAGGCGCATCAAGAAAGCGGCTCCAGACCAAGTCAACGAGGCGGCACT 1289
QY 403 GCGCAGCTGAGAAAGCCCTACCGGCGCTGAGTTCTGCACTAGCGGCGCTGATG 462
DB 1290 GCGGTCAACGTCAACGTCAACGACGCGCGCGCTGTCAGGATGCTGCGACAGAC 1349
QY 463 TGGATGCGCACTCAAGTCTGCGTCAAGAGCAAGACGCTACCGCAAGCGCGCGCTG 522
DB 1350 GCGTCTCGCGCGCGCGGCGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTG 1409
QY 523 TGTGACGTGCGGCGCGCGCGCTGCGCGCTGCTCAACGAGCGGCGGCGCGGATGAA 582
DB 1410 ATCCACTGCGGCGAGCTCGCGCTGCAAGGCGCAACAGGCTCCGCACTCAACGAG 1469
QY 583 GTGGGATGCGCAACCAATGCGATGTCGCGGAGATGCGCAATCTTCCCTTCTG 642
DB 1470 GCGGAGCGCGTGAAGCGCTATGCGCTGAAC-----GTGATGACCGCGCGCGCGCG 1524
QY 643 GAGCTGATGGAACACTGGAAGCTGTTTCAAGTGGGATCAACAGCGGCGGCGCAAT 702
DB 1525 GAGCGGAGTGCACAGTGGCGGTATGATGATCTCTCATACGCGCAACCTACGAG 1584
QY 703 CCGGTCAACCAACCGCATGTCATTCGCGCAATCTTCCGCTCAACCTTCCGATG 762
DB 1585 CTGGCAGCGGCGCTGCTGCGCTCGATTCGCGCGCTCTTCCAGAGCGCTTGAAGGCGAC 1644
QY 763 ATCTTGGCTACTACACGCGCTGAGCGCAAGCTGTTCTGCGACATGTCATGACGC 822
DB 1645 CTGAGCGAGAGAGACTGGGAGCAGCGCGCTCGCGCGCTCAAGCGTACGCGAGACGC 1704
QY 823 AGCCTGCAATCTGGGAGAGAAAGTGGCGCGTGCATTCGCGCGGCGCTGCA 872
DB 1705 GCGCGGCGAGTGTGCGCAGCGGCTGCTCATGCGCGCGTGTGCTCA 1754

```

RESULT 11

US-09-475-515-54

Sequence 54, Application US/09475515A

Patent No. 6602705

GENERAL INFORMATION:

APPLICANT: BARNETT, Susan

APPLICANT: ZUR MEGEDE, Jan

APPLICANT: SRIVASTAVA, Indresh

APPLICANT: LIAN, Ying

APPLICANT: HARTOG, Karin

APPLICANT: LIU, Hong

APPLICANT: GREER, Catherine

APPLICANT: SELBY, Mark

APPLICANT: WALKER, Christopher

TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
 OF VIRUS-LIKE PARTICLES
 FILE REFERENCE: 1621.002
 CURRENT APPLICATION NUMBER: US/09/475,515A
 CURRENT FILING DATE: 1999-12-30
 NUMBER OF SEQ ID NOS: 90
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 54

LENGTH: 1599
 TYPE: DNA
 ORGANISM: Artificial Sequence
 FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: gp120.modus4

US-09-475-515-54

Query Match 5.0%; Score 60.6; DB 4; Length 1599;
 Best Local Similarity 42.6%; Pred. No. 0.0014;
 Matches 434; Conservative 0; Mismatches 579; Indels 6; Gaps 2;

```

QY 196 CCGAATAGCGATGTCATGCAACCAACGCGCAACGATTTCCGCGCGCATCGAC 255
DB 364 CCGTGTGCTGACCTGTAACGCAACGCAAGCTGACCGGCAACCAACGCGCAAC 423
QY 256 GCGCGCAGCGCTGCGCGCGCGAGCTTCCGCGCAACATCACTTACCGAGCTGGCGCG 315
DB 424 AGCACACGCGGCAACCAACGACGCGGCAACGACGACGACGACGACGACGACGACGAC 483
QY 316 GACAAATTTATGCGCGCGCTGCGCGCT---GACCAAGGCGCGCAAGCGCATG 372
DB 484 AGCTGGAGAAAGATCCCGAGGCGGCAAGATCAAGATCAAGCTTCAACATCAACGAC 543
QY 373 GAGTTGACCAAGTCAATCTGACTTCCGCGCGCGCAAGTCAAGAGCGCTTACCGGCGCTG 432
DB 544 GTGGCGCAAGGTCAGAGAGATGACGCTGTTTCAAGCTGAGAGTGTGCTTCCATC 603
QY 433 GAGTTGTCGACATCAAGCGCGCGCTGCTGATGTCGCGCAACATCAAGTGTGTCAGAG 492
DB 604 GACAAACGAAAGCGAGCTTACGCTGATCAAGCTGCAACGAGGATGATCAACGAGGCG 663
QY 493 CAGAAAGTATCCGGAAGGCGCGCGCGCTGCTGATGCTGCGCGCGCGCGCTGCGGCT 552
DB 664 TGCCCAAGGATGAGTTGAGACCCATCCCATCTCTGCGCGCGCGCGCGCGCTTCCGCG 723
QY 553 GCGATCAAGCGCGCGCGCGCGCGAGATGAGTGGGATTCGCAACCAACGATGATGATC 612
DB 724 ATCTGAAAGTCAAGAGCAAGAAATTCAGAGCAACGCGCGCTGCAAGAGCTGAGCAC 783
QY 613 GCGGAGATCGCAATATGTTCCCTTCCGTTGAGAGTGAACAACCTGACCTGTTCCAG 672
DB 784 GTGCAATGACCAACGAGATCGCGCGCTGAGAGACCGAGCTGCTGAGACGCGAC 843
QY 673 TCGGCGATCAACACCGAGCGCGCGCAACATCCGCTGACCAACCGCATGTCATTCGCG 732
DB 844 CTGGCGGAGAGAGATGTCGCTGCGCTCCGAAACTTCAACGAAAGCGCAAGCACTATC 903
QY 733 GACATCTTTTGGCTCAACCTTCCGATGATCTTCCGCTACTACACCGCGTGAAGCGC 792
DB 904 ATCTGCAAGTGAACGATCTGAGAGATCACTGATCGCGCGCGCAACCAACGCGCT 963
QY 793 ACGGTTTTCGCAACATGTCATGACGCGCGCTGACATCTGAGAGAGAAAGTGGCGC 852
DB 964 AAGAGATCAACATGCGCGCGCGCGCGCGCTTCAAGCGCAACGAGATCACTGCGGAC 1023
QY 853 GTGATGCGCGCGCGCTGAGCTGATCAAGCGCGCGCGCTGCAAGAGATCGCATC 912
DB 1024 ATCGCGCAGGCGCACTGCAATCAAGAGCGCAACCTGAGCAACCTCGAGGAGATC 1083
QY 913 GAGCTCAACGAGATGACGCGAGTGGAGCTGCTGAAATGATCCGCTCTTGGCTATGCG 972
DB 1084 G---TGAGAGAGCTGCGGAGAGATTCGCGCAACCAAGACATCATCTTCAACGAGCAGC 1140
QY 973 CACTCTTCCGCGCTGCTGCTGCTACTACGATGCGGAGCGCGCGCTGAGAGCTGCGGAG 1032

```


Db 1141 AGCGGCGGAGACCCGAGATCGTGTTCACAGCTTCACTGCGGCGGAGATTCTTCTAC 1200
Qy 1033 GACATCGACACGAGCTGAAGCCCGGATGAGTGTCTCCATGAGCCGATGATGCTG 1092
Db 1201 TCGAACACGAGCAGCTGTTCACAGCAGCTGGAACATCACCGAGGAGTGAACAAGACC 1260
Qy 1093 CCGAGAGGAGATCCCGGCTGCGGCGCTATCCGAGACGACATCTGTATCGTGGGAG 1152
Db 1261 AAGGAGAACGACACATCATCTCTGCGCTCCGATCCGACATCATCAATGTCGCG 1320
Qy 1153 GACGTCGCGAGACATCAACCGGCTTCCGTCGTCGCGGAAACAACATCATCCGAA 1211
Db 1321 GAGGTGGGAGGCGCATGTACGCCCCCCCCCATCCGCGGCGACATCAAGTCAGCAGCAA 1379

RESULT 12
US-09-475-515-56
Sequence 56, Application US/09475515A

Patent No. 6602705
GENERAL INFORMATION:
APPLICANT: BARNETT, Susan
APPLICANT: ZUR MEGEDE, Jan
APPLICANT: SRIVASTAVA, Indresh
APPLICANT: LIAN, Ying
APPLICANT: HARTOG, Karin
APPLICANT: LIU, Hong
APPLICANT: GREER, Catherine
APPLICANT: SELBY, Mark
APPLICANT: WALKER, Christopher
TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
FILE REFERENCE: 1621.002
CURRENT APPLICATION NUMBER: US/09/475,515A
CURRENT FILING DATE: 1999-12-30
NUMBER OF SEQ ID NOS: 90
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 56
LENGTH: 2112
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: gp140.modus4
US-09-475-515-56

Query Match 5.0%; Score 60.6; DB 4; Length 2112;
Best Local Similarity 42.6%; Pred. No. 0.0015;
Matches 434; Conservative 0; Mismatches 579; Indels 6; Gaps 2;

Qy 196 CGCAAGTATGGGATGTATCGACCAACAACGCGACATTTCCGCGCGCATCGAC 255
Db 364 CCCCTGTGTGATACCTGTGATGCAACGACGCTGACCGGCGACCAACGCGACCAAC 423
Qy 256 GCGGCGCAGCCCTGCGCGCGGAGCTTCCGCGGACCAACATCATCAACGAGTGGCGCGC 315
Db 424 AGCACCAGGCGGACCAACGACGACGCGGACCAACGACGACCAACGACGACGAC 483
Qy 316 GACCAATTTCTATCGCGCGCTGCGCGACGCT--GACCAAGGCGCGCAAGCGCATCGGCATC 372
Db 484 AGCTGGAGAAAGATGCGCGGCGGAGATCAAGAACTGCAAGCTTCAATCAACACACAGC 543
Qy 373 GAGTTGACACATGTCATCTGCACTTCCGCGCGGAGCTGAGAGAAAGCTTACCGGGCGTC 432
Db 544 GTGCGCGAAGAGTGAAGAGATGACGCTGTCTTAACGCTGACGTCGTGCGCATC 603
Qy 433 GAGTTGTCGATCATCAACGCGCTGATGTGATGCGACCATCAAGTCGCTCGAAGAG 492
Db 604 GACCAAGACAAAGCGCAGCTACCGCTGATCACTGGAACAACAGCTGATCAACGAGCC 663
Qy 493 CAGAAAGTATCCGCGAGGCGCGCGCTGTGTGACGTGCGCGCGCGCGCTGCGCGCT 552
Db 664 TCGCCCAAGGTAGCTTCAAGCCCATCCCATCACTACCTGCGCGCGCGCTTCCGCGC 723
Qy 553 GCGATCAAGGCGCGCGCTGCGCGAGATGAAGTGGCATGCGCACCAACCAATGATGATC 612

Db 724 ATCTGAAGTGCAGAGAACAAAGATTTCACGCGACCGGCGCTTGCAGAAAGCTGAGCACC 783
Qy 613 CGCGAGATCGCAAAATCGTCCCTTCTGTGAGCTGATGAGACCTGAGCTGTTCAG 672
Db 784 GTGAGTGCACCAACGCGATCCGCGCTGTGTGAGCACCAGCTGCTGCTGAAACGCGACG 843
Qy 673 TCGGCGATCAACCGAGCGGCGGACCAATCCGCTACCAACCGCATGTCGATTCGCGC 732
Db 844 CTGCGCGAGAGAGATGTGTCTGTGCTCCGAGAACTTACCGACCAAGCGCAAGACATC 903
Qy 723 GACATCTTGTGCTCAACACCTTCCGATGATCTTCCGCTACTACACCGCGCTGAGCGC 792
Db 904 ATCTGACGCTGACAGATCGCTGTGAGATCACTGATCATCCGCCCAACAACAGCGCT 963
Qy 793 AGCTGTCTCGGACCATGTGATGATGAGCGCGAGCTTCAATCTGAGAGAGAAAGCTGCGC 852
Db 964 AAGAGCATTCACATGCGCGCGCGCGCTTCTTACGCGCACCGCGGACATCATCGCGGAC 1023
Qy 853 GTGCAATCCCGCGGCTGAGCTGATCAAGCCCGGCGCGCGCTGCAAGAGCATTCGCCATC 912
Db 1024 ATCCGCGAGGCGCCACTGCAACATGCAAGGCGCAATGCAACCAACCTTCAAGCATC 1083
Qy 913 GAGCTCAACGAGATGACCGGAGTGGAGCTGTGAAGTACCGCTTCCCTGCGCTATGAGC 972
Db 1084 G---TGAGAGAGCTGCGGAGCAGCTTGGCAACAACAGACCATCATCTTCAACAGCAGC 1140
Qy 973 CACTCTTTCGCGCTGCTGTCGCTACTGACGCTGCGAGCGCGCGTGTGAGCTGCGCGAG 1032
Db 1141 AGCGGCGGCGACCCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
Qy 1033 GACATCGACACGAGCTGAAGCCCGGATGAGTGTCTCCATGAGCCGATGTGTGCTG 1092
Db 1201 TCGAACACGAGCAGCTGTTCACAGCAGCTGGAACATCACCGAGGAGTGAACAAGACC 1260
Qy 1093 CCGAGAGGAGATCCCGGCTGCGGCGCTATGCGGAGCAGCAGCATCTGATCTGCGGAG 1152
Db 1261 AAGGAGAACGACACATCATCTCTGCGCTCCGCGCATCCGCGGATCATCAAGTCGCGAG 1320
Qy 1153 GACGTCGCGAGACATCAACCGGCTTCCGTCGTCGCGGAAACAACATCATCCGAA 1211
Db 1321 GAGGTGGGAGGCGCATGTACGCCCCCCCCCATCCGCGGCGACATCAAGTCAGCAGCAA 1379

RESULT 13
US-09-475-515-57
Sequence 57, Application US/09475515A

Patent No. 6602705
GENERAL INFORMATION:
APPLICANT: BARNETT, Susan
APPLICANT: ZUR MEGEDE, Jan
APPLICANT: SRIVASTAVA, Indresh
APPLICANT: LIAN, Ying
APPLICANT: HARTOG, Karin
APPLICANT: LIU, Hong
APPLICANT: GREER, Catherine
APPLICANT: SELBY, Mark
APPLICANT: WALKER, Christopher
TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
TITLE OF INVENTION: OF VIRUS-LIKE PARTICLES
FILE REFERENCE: 1621.002
CURRENT APPLICATION NUMBER: US/09/475,515A
CURRENT FILING DATE: 1999-12-30
NUMBER OF SEQ ID NOS: 90
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 57
LENGTH: 2112
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: gp140.mut.modus4
US-09-475-515-57

Query Match 5.0%; Score 60.6; DB 4; Length 2112;
 Best Local Similarity 42.6%; Pred. No. 0.0015;
 Matches 434; Conservative 0; Mismatches 579; Indels 6; Gaps 2;

196 CGCAAGTACGGCATGTGATCGACCAACCAACGCGACGATTCGGCCGGGATGAC 255
 364 CCCCTGTGCGTGAACCTGAACTGACGACGACGACGACGACGACGACGACGAC 423
 256 GGGGGCCAGCCCTGGGCGGCGGAGTTGGGCAACATCACTACACGACTGGCGCC 315
 424 AGACACGCGGCGACCAACGACACGAGGCGACCAACGACGACGACGACGACGAC 483
 316 GACATTTTATGCGCGCGGCGGCGAGCT--GACGACGGGCGGCGGCGGCGGATC 372
 484 AGCTGGAGAGAGATGCCCGAGGCGAGATCAAGAACTGCACTTCAATCACTACAC 543
 373 GAGTTGACACGCTCAATCTGCACTTCCGCGCGACGCTCGAGGAGAGCCCTACG 432
 544 GTGCGGACAAAGGTGAGAGAGTACAGCTGTTTCAAGCTGAGCGTGGTCCATC 603
 433 GAGTTGTGACATCAAGCCGCTGATGATGAGCGACATCAAGTCCTCGAAGAG 492
 604 GACACGACCAACCGCAGCTACCGCTGATCACTGCAACACGAGCTGATCAACCC 663
 493 CAGAACTGATCGCGAAGGCGCGCGCTGTGAGTGGCGGCGCGCGCGCGCT 552
 664 TGCCCCAAGGTGAGCTTCAGCCCATTCCTCATCACTACGCGCGCGCGCTTGGC 723
 553 GCCATCAAGGCGCGCGTGCAGCATGAGTGGCGATCCGACCAACATGCGATGATC 612
 724 ATCTGAGAGTGAAGGACAAAGATTCAACGCGCGCGCGCGCGCGCGCGCGCG 783
 613 CGGAGATGCGCAATCTTCCCTTGTGAGCTGATGAGACCTGAGACCTGTTCCAG 672
 784 GTGAGATGACCCACGCGCATCCCGCGGTGTGAGGACCCAGCTGTGTTGAAACG 843
 673 TCGGGCATCAACGCGCGCGCGCAATTCGGGTCAACGCGCATGCGAATCGCG 732
 844 CTGGCGGAGAGAGATCTGTGCTGCGCTCGAGAACTTCAACGCAACGCGACATC 903
 733 GACATCTTCTGCTCAACCTTCCGATGATCTTGGGCTACCAACGCGCTGAGAG 792
 904 ATCTGACACTGAACGAGTCCGTGAGATCACTGATCCGCGCGCGCGCGCGCG 963
 793 AGCTGTTTGTGCAACATGTCAGTGAAGCGGCGCTGCACTTGGAGAGAAAGTGG 852
 964 AAGAGCATCACTCGCGCGCGCGCGCGCTTCTACGCGCACGCGGACATCATCG 1023
 853 GTGCAATCGCGCGCGGCTGAGTCAACGCGCGCGCGCGCGCTGCAAGACATCG 912
 1024 ATTCGCGAGGCGCGCATGCAACATCAACGCGCGCGCGCGCGCGCGCGCGCG 1083
 913 GAGCTCAACGAGATGATCGCGAGTGGAGCTTGAAGTACCGCTCTTGGGCTATG 972
 1084 G---TGGAGAGTGGCGGAGAGTTCGGCAACAAAGACATCACTTCAACAGCAG 1140
 973 CACTCTTGGCGGTGCTGCACTAATCGGTGCGAGGCGCGCGTGGAGCTGCGCG 1032
 1141 AGCGGCGGCGACCGCGAGATCGTGTTCACAGCTTCACTGGGCGCGCGGTTCT 1200
 1033 GACATGCAACGAGCTGGAAGCGCGGCGATGATGATCCATGAGCGGATGATGAT 1092
 1201 TGCACACACGCGAGCTGTTCAACAGCACTGGAACATCAACGAGAGGTGAAACA 1260
 1093 CGGAGGCGGATCGCGGCGCGGCGCTATCGAGACGAGACATCTGATCGCGGAG 1152
 1261 AAGGAGAGACACATCACTCTGCGCTTCCGCAATCCGCGAGTCAATCATGATG 1320
 1153 GACGCTGCGGAGACATCAACGCGCTTCCGTTGCGTTCGGAACAAACATCACTCG 1211
 1321 GAGGTGGGCGAGGCGCATGATACGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1379

RESULT 14
 US-09-475-515-58
 ; Sequence 58, Application US/09475515A
 ; Patent No. 6602705
 ; GENERAL INFORMATION:
 ; APPLICANT: BARNETT, Susan
 ; APPLICANT: ZUR MEGEDE, Jan
 ; APPLICANT: SRIVASTAVA, Indresh
 ; APPLICANT: LIAN, Ying
 ; APPLICANT: HARTOG, Karin
 ; APPLICANT: LIU, Hong
 ; APPLICANT: GREER, Catherine
 ; APPLICANT: SELBY, Mark
 ; APPLICANT: WALKER, Christopher
 ; TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
 ; FILE OF INVENTION: OF VIRUS-LIKE PARTICLES
 ; FILE REFERENCE: 1621.002
 ; CURRENT APPLICATION NUMBER: US/09/475,515A
 ; CURRENT FILING DATE: 1999-12-30
 ; NUMBER OF SEQ ID NOS: 90
 ; SOFTWARE: Patentin Ver. 2.0
 ; SEQ ID NO 58
 ; LENGTH: 2181
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: gp140TM.modUS4
 US-09-475-515-58

Query Match 5.0%; Score 60.6; DB 4; Length 2181;
 Best Local Similarity 42.6%; Pred. No. 0.0015;
 Matches 434; Conservative 0; Mismatches 579; Indels 6; Gaps 2;

196 CGCAAGTACGGCATGTGATCGACCAACCAACGCGACGATTCGGCCGGGATGAC 255
 364 CCCCTGTGCGTGAACCTGAACTGACGACGACGACGACGACGACGACGACGAC 423
 256 GGGGGCCAGCCCTGGGCGGCGGAGTTGGGCAACATCACTACACGACTGGCGCC 315
 424 AGCAGCAGGCGGACCAACGACACGAGGCGACCAACGACGACGACGACGACGAC 483
 316 GACATTTTATGCGCGCGGCGGCGAGCT--GACCAAGGCGGCGGCGGCGGATC 372
 484 AGCTGGAGAGAGATGCCCGAGGCGAGATCAAGAACTGCACTTCAATCACTACAC 543
 256 GGGGGCCAGCCCTGGGCGGCGGAGTTGGGCAACATCACTACACGACTGGCGCC 315
 424 AGCAGCAGGCGGACCAACGACACGAGGCGACCAACGACGACGACGACGACGAC 483
 373 GAGTTGACCAAGTCAATCTGCACTTCCGCGCGAGCTTGAAGAACCTTACCGGCG 432
 544 GTGCGGACAAAGTGAAGAGAGTACAGCTGTTTCAAGCTGAGCGTGGTCCATC 603
 433 GAGTTGTGACATCAAGCGGCGCTGATGATGAGCGGCGGCGGCGGCGGCGGCT 552
 604 GACACGACCAACGCGGAGCTACCGCTGATCACTGCAACACGCGTGAATCCGCG 663
 493 CAGAACTGATCGCGAAGGCGCGCGGCTGTGAGTGGCGGCGGCGGCGGCGGCT 552
 664 TGCCCCAAGGTGAGCTTCAGCCCATTCCTCATCACTACGCGCGCGCGCTTGGC 723
 553 GCCATCAAGGCGCGGCGTGCAGCATGAGTGGCGATCCGACCAACATGCGATGATC 612
 724 ATCTGAGAGTGAAGGACAAAGATTCAACGCGCGCGCGCGCGCGCGCGCGCG 783
 613 CGGAGATGCGCAATCTTCCCTTGTGAGCTGATGAGACATCTGAGCTGTTCCAG 672
 784 GTGAGTGAACCCACGCGATCCGCGCGGTGAGACACCACTGCTGTGAACGCGAC 843
 673 TCGGGCATCAACGCGAGGCGCGGCAATCCGCTCAACCAACGAGATGATGATGAT 732
 844 CTGGCGGAGAGAGATGATGCTGCTGCGCTCCGAACTTCAACGAGACGCGACATC 903
 733 GACATCTTCTGCTCAACCTTCCGATGATCTTGGGCTACCAACGCGCTGAGAG 792
 904 ATCTGAGAGTGAAGAGATCTGCTGAGATCACTGATCCGCGCGCGCGCGCGCG 963

```

Qy 793 AGCGTGTTCGAGCATTCATGACGCGCATTCGATCTGGAGAAAGCTGGCC 852
Db 964 AAGAGCATTCACATCGGCCCCCGCCGCTTCAAGCCACCGGCGACATCTGGCCAG 1023
Qy 853 GTCGATTCGCGCGGGGTGAGCTGATCAAGCGGGGCGCGGTGAGAGACATTCGCATC 912
Db 1024 ATCCGCGAGGCCCATGCAACATCAAGCAAGGCGCAATTGACCAACCTTCAGACATTC 1083
Qy 913 GAGCTCAACGAGATGTACCGGAGTGGAGCTGTGAAAGTACCGCTCTTCGCTATGAC 972
Db 1084 G---TGAGAAAGCTGCGCAGAGAGTTGGGCAACAAGACATATCTTAAACAGACG 1140
Qy 973 CACTCTTCGCGGTGCTGTGCACTAATGAGTGTGAGGCGCGGTGAGCTGGCGAG 1032
Db 1141 AGCGCGCGGACCGCGAGATCTGTTTCCACAGCTTCACTGCGCGCGCACTTCTTAC 1200
Qy 1033 GACATCGAACCGAGCTGAAGCCGCGCATGTGTCTCATGAGCGCATGTGATGTG 1092
Db 1201 TGCACACACGACGCTGTTCACAGACCTGGAAATATCACGAGAGGTGAACAGACC 1260
Qy 1093 CCGAGGGGATCCCGGATCCGCGCTATTCGAGACAGACATCTCTATCTGCGGAG 1152
Db 1261 AAGGAAAGACACATATCTCTGCGCTGCGGATCGGCAATCATCAATGTGGAG 1320
Qy 1153 GACGCTCCGAGAACATCAACGCGCTTCCGTTGGTCCGGAACAACATCATTCGCA 1211
Db 1321 GAGGTGGGCAAGGCCATGTATGCGCCCCCATCTCCGCGGCAAGTACAGGACGCA 1379

```

RESULT 15
US-09-475-515-64
Sequence 64, Application US/09475515A
Patent No. 6602705

GENERAL INFORMATION:
APPLICANT: BARNETT, Susan
APPLICANT: ZUR MEGEDE, Jan
APPLICANT: SRIVASTAVA, Indresh
APPLICANT: LIAN, Yung
APPLICANT: HARTOG, Karin
APPLICANT: LIU, Hong
APPLICANT: GREER, Catherine
APPLICANT: SELBY, Mark
TITLE OF INVENTION: IMPROVED EXPRESSION OF HIV POLYPEPTIDES AND PRODUCTION
FILE REFERENCE: 1621.002
CURRENT APPLICATION NUMBER: US/09/475,515A
NUMBER OF SEQ ID NOS: 90
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 64
LENGTH: 2634
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: gp160.modus4
US-09-475-515-64

Query Match 5 0%; Score 60.6; DB 4; Length 2634;
Best Local Similarity 42.6%; Pred. No. 0.0015;
Matches 434; Conservative 0; Mismatches 579; Indels 6; Gaps 2;

```

Qy 196 CCGAAGTACGCGATGCTCATTCGACCAACAGCGCCAGATTTCCGCGCGATCGAC 255
Db 364 CCGCTGTGTGATACCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 423
Qy 256 GCGCGGCAAGCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 315
Db 424 AGCAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 483
Qy 316 GACATTTTCTATTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 372

```

```

Db 484 AGCTGGAGAAAGTCCCGAGGCGGAGATCAAGAACTGACGTTCAACATCACACAGC 543
Qy 373 GAGTTGACACAGTAAATCTGACTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 432
Db 544 GTGCGCGCAAGGTGCAAGAGAGTACAGCTGTTCTTCAAGCTGGAAGTGTGCGCATC 603
Qy 433 GAGTTGCTGACATGACGAGCCCTCATGTGATGAGTGCACCATCAATGCTCGAAGAG 492
Db 604 GACAAAGCAAGCGCGCTTACCGCTTATCACTGCAACACAGGTATATACCCAGGCC 663
Qy 493 CAGAAGTGAATCCGGAAGGCGCGCGGTGTGTAAGTGTGAGGCGCGCGCGCGCGCGCT 552
Db 664 TGCCCAAGAGTGAAGTTTGAAGCCATCCCATTCATCTGAGCGCGCGCGCGCTTCCG 723
Qy 553 GGCATGAAGCGCGCGGTCCCGAGCATGAATGAGATGCGCACCAACCAATGCGATGATC 612
Db 724 ATCTGAAGTGCAGGACAGAAAGTTCAACGACACCGGCGCGCGCGCGCGCGCGCGCG 783
Qy 613 CGCGAGATCGCAATGTTCCCTTGTGAGGTGATGACACTGGAAGCTGTTCAG 672
Db 784 GTGCAATGACATCCAGCATTCGCGCGGTGTGAGACACCACTGTCTGTAAACGCGAGC 843
Qy 673 TCGGCGATCAACACCGAGCGCGCGCAATCGGTCAACCAACCGATGTGCAATCGCG 732
Db 844 CTGCGCGAGAGAGATGTGTGTGCTCCGAGAACTTCAACGACAGCCAAAGCAATC 903
Qy 733 GACATCTTTGTGCTCAACATCTTCCGATGATCTTGGGTACTTACACCGCGGTGAGCG 792
Db 904 ATCTGAGCTGAAAGATTCGTGTGAGTCACTGATCTGCGCCCAACCAACACGCGT 963
Qy 793 AGCGTGTTCGAGCATGTGATGACGCGACCTTCGACATCTGGAGAAAGCTGGCC 852
Db 964 AAGAGCATTCACATCGGCCCCCGCCGCTTCAAGCCACCGGAGACATCTGGCCAG 1023
Qy 853 GTCGATTCGCGCGGCTGAGCTGATCAAGCGCGCGCGCGCTGAGAGCATTCGCCATC 912
Db 1024 ATCCGCGAGGCCCATGCAACATCAAGAGCCCACTGAGACCAACCTTCAGAGATC 1083
Qy 913 GAGCTCAACGAGATGTACCGGAGTGGAGCTGTGAAAGTACCGCTCTTCGCGTATGCG 972
Db 1084 G---TGAGAAAGCTGCGGAGAGTTCGCAACCAAGACCATCATCTTCAACAGACG 1140
Qy 973 CACTCTTCGCGGTGCTGTGCACTAATGAGTGTGAGGCGCGGTGAGAGCTGGCGAG 1032
Db 1141 AGCGCGCGGACCGCGAGTGTGTTCCACAGCTTCAACTGCGCGCGAGTTCTTAC 1200
Qy 1033 GACATCGAACCGAGCTGAAGCCGCGCATGTGTCTCATGAGCCGATGTGATGCTG 1092
Db 1201 TGCACACACGACGCTGTTCACAGACCTGGAAATATCACGAGAGGTGAACAGACC 1260
Qy 1093 CCGAGGGGATCCCGGATCCGCGCTATTCGCGAGACAGACATCTCTATCTGCGGAG 1152
Db 1261 AAGGAAAGACACCATATCTGCGCTGCGGATCGGCAATCATCAATGTGGAG 1320
Qy 1153 GACGCTCCGAGAACATCAACGCGCTTCCGTTGGTCCGGAACAACATCATTCGCA 1211
Db 1321 GAGGTGGGCAAGGCCATGTATGCGCCCCCATCTCCGCGGCAAGTACAGGACGCA 1379

```

Search completed: July 7, 2005, 23:34:22
Job time : 240 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 7, 2005, 22:16:36 ; Search time 828 Seconds
(without alignments)
9189.996 Million cell updates/sec

Title: US-10-807-228a-2

Perfect score: 1212
Sequence: 1 ATGACTGACGACATGTGCA.....AACACACATATCCGCAAC 1212

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 6330943 seqs, 3139157217 residues

Total number of hits satisfying chosen parameters: 12661886

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA.*

```
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	984.8	81.3	1215	US-10-251-078-1	Sequence 1, Appl
2	984.8	81.3	1215	US-10-251-078-17	Sequence 17, Appl
3	981.6	81.0	1215	US-10-251-078-9	Sequence 9, Appl
4	980	80.9	1212	US-10-251-078-11	Sequence 11, Appl
5	980	80.9	1212	US-10-251-078-13	Sequence 13, Appl
6	980	80.9	1215	US-10-251-078-15	Sequence 15, Appl
7	978.4	80.7	1212	US-10-251-078-21	Sequence 21, Appl

8	976.8	80.6	1212	US-10-251-078-19	Sequence 19, Appl
9	73.8	6.1	2733	US-10-437-963-88003	Sequence 88003, A
10	71.8	5.9	1262	US-10-428-115-109294	Sequence 109294, A
11	69.8	5.8	1483	US-10-437-963-38426	Sequence 38426, A
12	69.6	5.7	1212	US-10-411-910A-222	Sequence 222, App
13	68.6	5.7	1169	US-10-437-963-57011	Sequence 57011, A
14	68	5.6	834	US-10-156-761-4932	Sequence 4932, Ap
15	68	5.6	985	US-10-363-345A-31407	Sequence 31407, A
16	68	5.6	985	US-10-363-345A-31408	Sequence 31408, A
17	68	5.6	985	US-10-363-483A-31407	Sequence 31407, A
18	68	5.6	985	US-10-363-483A-31408	Sequence 31408, A
19	68	5.6	9025608	US-10-156-761-1	Sequence 1, Appl
20	67.8	5.6	1227	US-10-399-693-15	Sequence 15, Appl
21	67.6	5.6	1566	US-10-817-483-10	Sequence 30, Appl
22	67.2	5.5	65140	US-10-203-295-1	Sequence 1, Appl
23	67.2	5.5	125401	US-10-203-295-35	Sequence 35, Appl
24	66.6	5.5	2052	US-10-282-122A-11731	Sequence 11731, A
25	66	5.4	2960	US-10-437-963-29365	Sequence 29365, A
26	65.6	5.4	1516	US-10-437-963-29365	Sequence 29365, A
27	65.4	5.4	2569	US-10-437-963-88004	Sequence 88004, A
28	65.4	5.4	4512	US-10-156-761-2222	Sequence 2222, Ap
29	65.4	5.4	138203	US-10-819-386A-1	Sequence 1, Appl
30	65.4	5.4	9025608	US-10-156-761-1	Sequence 1, Appl
31	65.2	5.4	2712	US-10-282-122A-24663	Sequence 24663, A
32	65	5.4	573	US-10-156-761-1931	Sequence 3931, Ap
33	64.6	5.3	9546	US-10-463-074-2	Sequence 2, Appl
34	64.4	5.3	2037	US-10-228-063-48	Sequence 48, Appl
35	64	5.3	1053	US-10-156-761-5241	Sequence 5241, Ap
36	64	5.3	1294	US-09-748-033-2	Sequence 2, Appl
37	63.6	5.2	690	US-10-006-922-27	Sequence 27, Appl
38	63.6	5.2	690	US-10-081-864-19	Sequence 19, Appl
39	63.4	5.2	1161	US-10-282-122A-11406	Sequence 11406, A
40	63.4	5.2	2026	US-10-437-963-12278	Sequence 32278, A
41	63.4	5.2	2649	US-10-437-963-112	Sequence 112, App
42	63.2	5.2	9785	US-10-359-120-176	Sequence 176, App
43	63	5.2	1752	US-10-411-910A-234	Sequence 234, App
44	63	5.2	14335	US-10-156-761-412	Sequence 412, App
45	63	5.2	100000	US-10-156-761-15103	Sequence 15103, A

ALIGNMENTS

```
RESULT 1
US-10-251-078-1
; Sequence 1, Application US/10251078
; Publication No. US20030119084A1
; GENERAL INFORMATION:
; APPLICANT: Roche Diagnostics GmbH
; APPLICANT: Shao, Zhixin
; APPLICANT: Schmuck, Rainer
; APPLICANT: Kratzsch, Peter
; APPLICANT: Kemlins, Janet
; APPLICANT: Weisner, Harald
; TITLE OF INVENTION: Variants of an Erwinia-type creatinase
; FILE REFERENCE: 20981 EP
; CURRENT APPLICATION NUMBER: US/10/251,078
; CURRENT FILING DATE: 2002-09-20
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1215
; TYPE: DNA
; ORGANISM: Erwinia sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1215)
; US-10-251-078-1
Query Match      81.3%; Score 984.8; DB 15; Length 1215;
Best Local Similarity 80.3%; Pred. No. 1.2e+258;
Matches 1070; Conservative 0; Mismatches 142; Indels 0;
```

```
QY 1 ATGACTGAGCAATGTTGACGCGTGAATGAAATGCGCAACAGCGGAGAAAGATTATTCGCGG 60
DB 1 ATGACTGAGCAATGTTGACGCGTGAATGAAATGCGCAACAGCGGAGAAAGATTATTCGCGG 60
QY 61 TTTTGGAGATGCGAGATGACCGCGCGCAAAAAGAGTTGCGGCGTGAATGCGCAAGAAC 120
DB 61 TTTTGGAGATGCGAGATGACCGCGCGCAAAAAGAGTTGCGGCGTGAATGCGCAAGAAC 120
QY 121 AATGTCGATGCGGCGCTGTTCACTCTTATCACTGATCACTAATTCCGCTGGCTG 180
DB 121 GACGTGACGCTGCGCTGTTCACTCTTATCACTGATCACTAATTCCGCTGGCTG 180
QY 181 TACTGCTATTTGCGACGCAAGTACGCGCATGCTATGACCAACAAGCGCAAGCAAT 240
DB 181 TACTGCTATTTGCGACGCAAGTACGCGCATGCTATGACCAACAAGCGCAAGCAAT 240
QY 241 TCGGCGCGGATGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
DB 241 TCGGCGCGGATGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
QY 301 ACCGACTGCGCGCGCGCAATTTCTATCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
DB 301 ACCGACTGCGCGCGCGCAATTTCTATCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
QY 361 CGCATGCGGATGAGTTGCAACGCTCAATCTTCACTTCCGCGCGCGCGCGCGCGCG 420
DB 361 CGCATGCGGATGAGTTGCAACGCTCAATCTTCACTTCCGCGCGCGCGCGCGCGCG 420
QY 421 CTACGCGGCGTGAATGCTGATCGCGCAATCAAGCGCGCGCGCGCGCGCGCGCGCG 480
DB 421 CTACGCGGCGTGAATGCTGATCGCGCAATCAAGCGCGCGCGCGCGCGCGCGCGCG 480
QY 481 TCGCTGGAAGAGCAAGCTGATCGCGCAAGCGCGCGCGCGCGCGCGCGCGCGCG 540
DB 481 TCGCTGGAAGAGCAAGCTGATCGCGCAAGCGCGCGCGCGCGCGCGCGCGCGCG 540
QY 541 GCGTGGGCGGCTGCGCATCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
DB 541 GCGTGGGCGGCTGCGCATCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
QY 601 AATGCGATGATCGCGGAGATCGCGCAATGTTCCCTTCTGTTGAGCTGATGGAAC 660
DB 601 AATGCGATGATCGCGGAGATCGCGCAATGTTCCCTTCTGTTGAGCTGATGGAAC 660
QY 661 ACTGCGTTCAGTGGGCGATCAACACCGGCGCGCGCGCGCGCGCGCGCGCGCG 720
DB 661 ACTGCGTTCAGTGGGCGATCAACACCGGCGCGCGCGCGCGCGCGCGCGCGCG 720
QY 721 GTGCAATCGGCGGACATCTTTCGCTCAACCTTCCGATGATCTTCCGCTACTAC 780
DB 721 GTGCAATCGGCGGACATCTTTCGCTCAACCTTCCGATGATCTTCCGCTACTAC 780
QY 781 GCGCTGAGAGCGACGCTGTTCTGCGACCAATGTCATGACGCGCGCGCGCGCGCG 840
DB 781 GCGCTGAGAGCGACGCTGTTCTGCGACCAATGTCATGACGCGCGCGCGCGCGCG 840
QY 841 AAGAAAGTGGCGCTGATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
DB 841 AAGAAAGTGGCGCTGATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
QY 901 GACATGCGCATGAGCTCAACGAGATGTAACCGGAGATGGGACTGTAAGTACCGCT 960
DB 901 GACATGCGCATGAGCTCAACGAGATGTAACCGGAGATGGGACTGTAAGTACCGCT 960
QY 961 TTTGCGCTATGAGCACTCTTTCGCGCTGCTGCTCACTACGCGTCCGAGGCGCG 1020
DB 961 TTTGCGCTATGAGCACTCTTTCGCGCTGCTGCTCACTACGCGTCCGAGGCGCG 1020
QY 1021 GAGCTGCGGAGAGCAATGCAACCGAGCTGAAGCGCGCGCGCGCGCGCGCGCG 1080
DB 1021 GAGCTGCGGAGAGCAATGCAACCGAGCTGAAGCGCGCGCGCGCGCGCGCGCG 1080
QY 1081 ATGCTGATGCTGCGGAGGAGCAATGCGCGGCGCGCGCGCGCGCGCGCGCGCG 1140
```

```
DB 1081 ATGCTGATGCTGCGGAGGAGCAATGCGCGGCGCGCGCGCGCGCGCGCGCGCG 1140
QY 1141 ATGCTGCGGAGAGAGCGGCTGCGGAGAGCAATCAACCGCGTCCGCTGCTCGGAA 1200
DB 1141 ATGCTGCGGAGAGAGCGGCGCGGAGAGCAATCAACCGCGTCCGCTGCTCGGAA 1200
QY 1201 ATCATCGGCAAC 1212
DB 1201 ATCATCGGCAAC 1212

RESULT 2
US-10-251-078-17
; Sequence 17, Application US/10251078
; Publication No. US20030119084A1
; GENERAL INFORMATION:
; APPLICANT: Roche Diagnostics GmbH
; APPLICANT: Shao, Zhixin
; APPLICANT: Schmuck, Rainer
; APPLICANT: Kratzsch, Peter
; APPLICANT: Kenkies, Janet
; APPLICANT: Welser, Harald
; TITLE OF INVENTION: Variants of an Erwinia-type creatinase
; FILE REFERENCE: 20981 EP
; CURRENT APPLICATION NUMBER: US/10/251,078
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 1215
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:variant CTrg2
; NAME/KEY: CDS
; LOCATION: (1)..(1215)
US-10-251-078-17

Query Match 81.3%; Score 984.8; DB 15; Length 1215;
Best Local Similarity 88.3%; Pred. No. 1.2e-258;
Matches 1070; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

QY 1 ATGACTGAGCAATGTTGACGCGTGAATGAAATGCGCAACAGCGGAGAAAGATTATTCGCGG 60
DB 1 ATGACTGAGCAATGTTGACGCGTGAATGAAATGCGCAACAGCGGAGAAAGATTATTCGCGG 60
QY 61 TTTTGGAGATGCGAGATGACCGCGCGCAAAAAGAGTTGCGGCGTGAATGCGCAAGAAC 120
DB 61 TTTTGGAGATGCGAGATGACCGCGCGCAAAAAGAGTTGCGGCGTGAATGCGCAAGAAC 120
QY 121 AATGTCGATGCGGCGCTGTTCACTCTTATCACTGATCACTAATTCCGCTGGCTG 180
DB 121 GACGTGACGCTGCGCTGTTCACTCTTATCACTGATCACTAATTCCGCTGGCTG 180
QY 181 TACTGCTATTTGCGACGCAAGTACGCGCATGCTATGACCAACAAGCGCAAGCAAT 240
DB 181 TACTGCTATTTGCGACGCAAGTACGCGCATGCTATGACCAACAAGCGCAAGCAAT 240
QY 241 TCGGCGCGGATGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
DB 241 TCGGCGCGGATGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
QY 301 ACCGACTGCGCGCGCGCAATTTCTATCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
DB 301 ACCGACTGCGCGCGCGCAATTTCTATCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
QY 361 CGCATGCGGATGAGTTGCAACGCTCAATCTTCACTTCCGCGCGCGCGCGCGCGCG 420
DB 361 CGCATGCGGATGAGTTGCAACGCTCAATCTTCACTTCCGCGCGCGCGCGCGCGCG 420
QY 421 CTACGCGGCGTGAATGCTGATCGCGCAATCAAGCGCGCGCGCGCGCGCGCGCGCG 480
```

Db	421	TTGCTCCGCGGTGAGTTCCTGTCATATTCGCTCAACCTGTCATGTGGATTCGCCACGCTCAAG		480
Qy	481	TGGCTCGAAGAGCAGAAAGCTGATTCGCGGAAAGCGCCGCTGTGTGAACGTCCGCGGCGCG		540
Db	481	TGCGTCGAAGAGCAGAAAGCTGATTCGCCGAGGGGTGCCGTATCTGGCAACGTCGCGGTGCC		540
Qy	541	GCCTGGCGCGCTGCCATCAAGGCGCGCTGCCGAGCATGAAGTGGCGATTCGCCACACC		600
Db	541	GCCGTGGTTCGCGCCCTCAAGGCGCGCTGCCGAGCATGAAGTGGCGATTCGCCACACC		600
Qy	601	AATGCGATGATCCGCGAGATTCGCCAATGCTGCCCTTGTTGGAGCTGATGACACTGG		660
Db	601	AATGCGATGATCCGCGAGATTCGCCAATGCTGCCCTTGTTGGAGCTGATGACACTGG		660
Qy	661	ACCTGGTTCAGTCCGGGCATCAACCGGACCGGCGCACAAATCCGCTCAACCAACCGCATC		720
Db	661	ACCTGGTTCAGTCCGGGCATCAACCGGACCGGCGCACAAATCCGCTCAACCAACCGCATC		720
Qy	721	GTGCAATCCGCGGCATCTCTTTGCTCAACACTTCCCGATGATCTTCCGCTACACCC		780
Db	721	GTGCAATCCGCGGCATCTCTTTGCTCAACACTTCCCGATGATCTTCCGCTACACCG		780
Qy	781	GCGCTGGAGCGGACGCTGTTCTTGCGACAAATGTGATGAGCGGCGCTCGACATCTGGGAG		840
Db	781	GCGCTGGAGCGGACGCTGTTCTTGCGACAAATGTGATGAGCGGCGCTCGACATCTGGGAG		840
Qy	841	AAAGACGTGGCCGTGCATCGCGCGCGGCTCGAGCTGATCAAGCCGCGCGCGCTGCAAG		900
Db	841	AAAGACGTGGCCGTGCATCGCGCGCGGCTCGAGCTGATCAAGCCGCGCGCGCTGCAAG		900
Qy	901	GACATTCGCCATCGAGCTCAACGAGATGACCCGAGATGGGACCTGCTGAAGTACCGCTCC		960
Db	901	GATATTCGCCATCGAATCTCAACGAGATGACCCGAGATGGGATCTGCTGAAGTACCGCTCC		960
Qy	961	TTCCGCTATAGGCACTCCCTCCGCGGTGCTGTCATCTAGTGGTGGGAGGCGCGCGTG		1020
Db	961	TTCCGCTATAGGCACTCCCTCCGCGGTGCTGTCATCTAGTGGTGGGAGGCGCGCGTG		1020
Qy	1021	GAGCTGCGCGAGACATCGACACCGAGCTGAAGCCCGGCGATGGTGTCTCAATGAGCCG		1080
Db	1021	GAGCTGCGCGAGACATCGACATCCGATACCGGCTGCGAGCCCGGATGGTGTCTCAATGAGCCG		1080
Qy	1081	ATGTGATGCTCCGCGAGGGGACATGCCCGGTGCCGCGGCTATTCGCGAGACGACATCTTG		1140
Db	1081	ATGTGATGCTCCGCGAGGGGACATGCCCGGTGCCGCGGCTATTCGCGAGACGACATCTTG		1140
Qy	1141	ATCGTGGGGGAGAGCGGTGCCGAGAACATCACCGGCTTCCGCTGGTTCGGAACACCAAC		1200
Db	1141	ATCGTGGGGGAGAGCGGTGCCGAGAACATCACCGGCTTCCGCTGGTTCGGAACACCAAC		1200
Qy	1201	ATCATCCGCAAC 1212		
Db	1201	ATCATCCGCAAC 1212		

RESULT 3
 US-10-251-078-9
 ; Sequence 9, Application US/10251078
 ; Publication No. US20030119084A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Roche Diagnostics GmbH
 ; APPLICANT: Shao, Zhixin
 ; APPLICANT: Schmuick, Rainer
 ; APPLICANT: Kretzsch, Peter
 ; APPLICANT: Kenkies, Janet
 ; APPLICANT: Weisner, Harald
 ; TITLE OF INVENTION: Variants of an Erwinia-type creatinase
 ; FILE REFERENCE: 20981 EP
 ; CURRENT APPLICATION NUMBER: US/10/251,078
 ; CURRENT FILING DATE: 2002-09-20
 ; NUMBER OF SEQ ID NOS: 32
 ; SOFTWARE: PatentIn version 3.1

```

/ SEQ ID NO 9
/ LENGTH: 1215
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence:variant CT1m24
/ NAME/KEY: CDS
/ LOCATION: (1)..(1215)
US-10-251-078-9

Query Match      81.0%; Score 981.6; DB 15; Length 1215;
Best Local Similarity 88.1%; Pred. No. 9e-258;
Matches 1068; Conservative 0; Mismatches 144; Indels 0; Gaps 0

OY      1 ATGACTGACGACATGTTGCACTGTATGAAATGCGACAAACGCGCAGAAAGATTATTCGCG 60
DB      1 ATGACTGACGACATGTTGCACTGTATGAAATGCGACAAATGCTGAGAGAAATATTCGCC 60

OY      61 TTTTCGAGATCCGAGATGACCCGCCGCAAAACGACGTTTCGCGCTGATGCGCAGAAC 120
DB      61 TTTTCGAGATCCGAGATGACCGCGCCGCAAGTGAAGTGCAGCGCTGATGCGCGAAAC 120

OY      121 AATGTCGATGCGCGCTGTTCACTCTTATCACTGATCAACTATCTATTCGCGTGGCTG 180
DB      121 GACGTGACGCTGCGCTGTTCACTCTTATCACTGATCAACTATCTTGTGATTCCTG 180

OY      181 TACTGCTATTTTCGACGCGAATGACGCGCATGTCATGCAACCAACACGCGACGATTT 240
DB      181 TACTGCTATTTTCGCGCGCAAAATACGCGCATGTCATGCAACGACGACATGCGACATC 240

OY      241 TCGCGCGGCGATGACGCGCGCGCAAGCCCTGCGCGCGAGCTTCGCGCAACAATCACC 300
DB      241 TCGCGCGGCGATGATGCGCGCTGACGCCCTTGGCGCGTACGCTTCGCGCAACAATCACC 300

OY      301 ACCGACTGCGCGCGCGACAAATTTCTATCGCGCGTGGCGAGTGACACGCGCGCCAG 360
DB      301 ACCGACTGCGCGCGCGACAACTTTCTACAGCGCGCTGCGCAACTCACCCCGCGCCAG 360

OY      361 CGCATTCGCGCATGGAATTCGACCAACGCTCATCTTCGCGCGCGCGCATGCGAGAAC 420
DB      361 CGCATTCGGAATTCGATTCGATCACTGTGACCTTGAATTCGCGCGCAAGCTTCGAGAG 420

OY      421 CTACCGGCGCTGAGTTCGTGACATCAGCCGCTGATGTGATGCGCACATCAAG 480
DB      421 CTGCGCGGCGTGAATTCGTGATTCGATTCAGACGTCATGTGATGCGCACGCTCAAG 480

OY      481 TCGCTCGAAGACAGCAAGCTGATTCGCGCAAGGCGCGCGTGTGTGACGTTCGCGCG 540
DB      481 TCGCTCGAAGACAGCAAGCTGATTCGCGCAAGGCGTGCCTGATTCGCAACGTCGCGG 540

OY      541 GCGTCGCGGCGTGCATCAAGGCGCGGCGCGCGACGATGAAGTGGCGATTCGCGCAC 600
DB      541 GCGTCGCGTTCGCGCGCTCAAGGCGCGGCGTTCGCGAGACAGAGTTCGCGATTCGCGCAC 600

OY      601 AATGCGATGATTCGCGAGATCGCCAAATCGTTCCCTTCGTGAGCTGATGACACCTGG 660
DB      601 AATGCGATGATTCGCGAGATTCGCGAAATCGTTCCCTTCGTGGAATGATGACACCTGG 660

OY      661 ACCTGCTTCAGTCCGCGCATCAACACCGAGCGCGCGCAAAATCCGATGACCAACGCGATC 720
DB      661 ACCTGCTTCAGTCCGCGCATCAACACCGAGCGCGCGCAAAATCCGATGACCAACGCGATC 720

OY      721 GTGCAATCCGCGGACATCCTTTGCTGCAACACTTCCGATGATCTTGGGCTACTACAC 780
DB      721 GTGCAATCCGCGGACATCTCTGCTGCTCAACAGCTTCCGATGATCTTGGGCTACTACAG 780

OY      781 GCGCTGAGCGCACGCTGTTCTTCGACACCATGTGATGACGCGACGCTTCGACATCTGGAG 840
DB      781 GCGCTGAGCGCAAGCTGTTCTTCGATCCACCATGTGAGCGACGCTTGTACACTTGGAG 840

OY      841 AAGAACTGCGCGCTGATCGCGCGGCGCTGACCTGATTAAGCGCGGCGCGCTGCAAG 900
DB      841 AAGAACTGCGCGCTGATCGCGCGGCGCTGACCTGATTAAGCGCGGCGCGCTGCAAG 900

```



```

; Sequence 13, Application US/10251078
; Publication No. US20030119084A1
; GENERAL INFORMATION:
; APPLICANT: Roche Diagnostics GmbH
; APPLICANT: Shao, Zhixin
; APPLICANT: Schuck, Rainer
; APPLICANT: Kratzsch, Peter
; APPLICANT: Kenkies, Janet
; APPLICANT: Weisner, Harald
; TITLE OF INVENTION: Variants of an Erwinia-type creatinase
; FILE REFERENCE: 20981 EP
; CURRENT APPLICATION NUMBER: US/10/251,078
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 1212
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:variant CT2m10
; NAME/KEY: CDS
; LOCATION: (1)..(1212)
; US-10-251-078-13

Query Match      80.9%; Score 980; DB 15; Length 1212;
Best Local Similarity 88.0%; Pred. No. 2.5e-257;
Matches 1067; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

QY 1 ATGACTGACGACATGTTGACGATGATGAAATGACACAGCGGAGAAATTAATTCGCCG 60
DB 1 ATGACTGACGACATGTTGACGATGATGAAATGACACAGCGGAGAAATTAATTCGCCG 60
QY 61 TTTTGGATGCGGATGACCCGCCGCAAAAGACGTTGCGGCTGATGCGCAAGAAC 120
DB 61 TTTTCCGATGCGGATGACCGGCCGCAAGATGATGCGGCTGATGCGCAAGAAC 120
QY 121 AATGCGATGCGGCGGTGTTCACTCTTATCACTGATCACTACTATTCGGCTGCGG 180
DB 121 GAGCTGACGCGCGGTGTTCACTCTTATCACTGATCACTACTATTCGGCTGCGG 180
QY 181 TACTGCTATTTTCGACGCAAGTACGATGATGATGACACACAGCAAGCAAGAT 240
DB 181 TACTGCTATTTTCGACGCAAGTACGATGATGATGACACACAGCAAGCAAGAT 240
QY 241 TCGGCGCGGATGACGCGCGGCAAGCCCTGCGCGCGGCAAGCTTTCGCGCAACTAC 300
DB 241 TCGGCGCGGATGACGCGCGGCAAGCCCTGCGCGCGGCAAGCTTTCGCGCAACTAC 300
QY 301 ACCGATGCGCGCGGCAAGTATTTCTATGCGCGCGGCTGCGCGCAAGTACCGCGCGCAAG 360
DB 301 ACCGATGCGCGCGGCAAGTATTTCTATGCGCGCGGCTGCGCGCAAGTACCGCGCGCAAG 360
QY 361 CGCATGCGGATGACGATGACACGATGATGACCTTTCGCGCGGCAAGTACCGCGCGCAAG 420
DB 361 CGCATGCGGATGACGATGACACGATGATGACCTTTCGCGCGGCAAGTACCGCGCGCAAG 420
QY 421 CTACCGCGGCGTTCGATGATGATGACACGCGCTGATGATGATGATGATGATGATGATGAT 480
DB 421 CTACCGCGGCGTTCGATGATGATGACACGCGCTGATGATGATGATGATGATGATGATGAT 480
QY 481 TCGCTCGAAGACGACAGTATTCGCGCAAGCGCGCGGCTGATGATGATGATGATGATGATGATGAT 540
DB 481 TCGCTCGAAGACGACAGTATTCGCGCAAGCGCGCGGCTGATGATGATGATGATGATGATGATGAT 540
QY 541 GCTTCGCGGCGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
DB 541 GCTTCGCGGCGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
QY 601 AATGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
DB 601 AATGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660

```

```

QY 661 ACTGCTTCAGTTCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
DB 661 ACTGCTTCAGTTCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
QY 721 GTGCAATTCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
DB 721 GTGCAATTCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
QY 781 GCGCTGAGACGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
DB 781 GCGCTGAGACGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
QY 841 AAGACGTCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
DB 841 AAGACGTCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
QY 901 GACATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
DB 901 GATATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
QY 961 TTTGCGCTATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
DB 961 TTTGCGCTATGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
QY 1021 GAGCTGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
DB 1021 GAGCTGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
QY 1081 ATGCTGATGCTGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
DB 1081 ATGCTGATGCTGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
QY 1141 ATGCTGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
DB 1141 ATGCTGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
QY 1201 ATCATTCGCGAAC 1212
DB 1201 ATCATTCGCGAAC 1212

RESULT 6
US-10-251-078-15
; Sequence 15, Application US/10251078
; Publication No. US20030119084A1
; GENERAL INFORMATION:
; APPLICANT: Roche Diagnostics GmbH
; APPLICANT: Shao, Zhixin
; APPLICANT: Schuck, Rainer
; APPLICANT: Kratzsch, Peter
; APPLICANT: Kenkies, Janet
; APPLICANT: Weisner, Harald
; TITLE OF INVENTION: Variants of an Erwinia-type creatinase
; FILE REFERENCE: 20981 EP
; CURRENT APPLICATION NUMBER: US/10/251,078
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 1215
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:variant CT2m28
; NAME/KEY: CDS
; LOCATION: (1)..(1215)
; US-10-251-078-15

Query Match      80.9%; Score 980; DB 15; Length 1215;
Best Local Similarity 88.0%; Pred. No. 2.5e-257;
Matches 1067; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

```

QY 1 ATGACTGACGACATGTTGCACTGATGAAATGGCAACAAGCGGAGAAATTAATGCGG 60
 DB 1 ATGACTGACGACATGTTGCACTGATGAAATGGCAACAAGCGGAGAAATTAATGCGG 60
 QY 61 TTTTCGGAATGCGAGATGACCGCGCGCAAAAAGAGCTTCCGCGTGTGATGCGCAAGAAC 120
 DB 61 TTTTCGGAATGCGAGATGACCGCGCGCAAAAAGAGCTTCCGCGTGTGATGCGCAAGAAC 120
 QY 121 AATGTGATGCGGCGCTGTTCACTCTTATCACTGATCAACTACTATTCGCGCTG 180
 DB 121 AATGTGATGCGGCGCTGTTCACTCTTATCACTGATCAACTACTATTCGCGCTG 180
 QY 181 TACTGCTATTTGCGAGCGCAAGTACGCGATGATGACCAACAACGCGACGACGATT 240
 DB 181 TACTGCTATTTGCGAGCGCAAGTACGCGATGATGACCAACAACGCGACGACGATT 240
 QY 241 TCGGCGGCGATGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
 DB 241 TCGGCGGCGATGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
 QY 301 ACCGACTGCGCGCGCGCGCAATTTCTATCGCGCGCGCGCGCGCGCGCGCGCGCG 360
 DB 301 ACCGACTGCGCGCGCGCGCAATTTCTATCGCGCGCGCGCGCGCGCGCGCGCGCG 360
 QY 361 CGGATGCGGATGCGAGTTGCAACGCGCACTGCGACTTCGCGCGCGCGCGCGCGCG 420
 DB 361 CGGATGCGGATGCGAGTTGCAACGCGCACTGCGACTTCGCGCGCGCGCGCGCGCG 420
 QY 421 CTACCGGCGGCTGCGAGTTGCGCATCAGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
 DB 421 CTACCGGCGGCTGCGAGTTGCGCATCAGCGCGCGCGCGCGCGCGCGCGCGCGCG 480
 QY 481 TCGCTCGAAGAGAGAGAGCTGATCGCGAGAGCGCGCGCGCGCGCGCGCGCGCG 540
 DB 481 TCGCTCGAAGAGAGAGAGCTGATCGCGAGAGCGCGCGCGCGCGCGCGCGCGCG 540
 QY 541 GCGTGGCGCGGCTGCGCATCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
 DB 541 GCGTGGCGGCTGCGCATCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
 QY 601 AATGCGATGATCGCGAGATGCGCAATCGTTCCCTTCGATGAGTGAACGCTG 660
 DB 601 AATGCGATGATCGCGAGATGCGCAATCGTTCCCTTCGATGAGTGAACGCTG 660
 QY 661 ACCTGCTTCAAGTCGCGCATCAACGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
 DB 661 ACCTGCTTCAAGTCGCGCATCAACGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
 QY 721 GTCGCAATCGCGCGCATTCCTTTGCTTCAACCTTCCCATGATTTTCGCTACTAC 780
 DB 721 GTCGCAATCGCGCGCATTCCTTTGCTTCAACCTTCCCATGATTTTCGCTACTAC 780
 QY 781 GCGCTGAGAGCGAGCTGTTCTGCGACATGTCGATGACGCGCGCGCGCGCGCGCG 840
 DB 781 GCGCTGAGAGCGAGCTGTTCTGCGACATGTCGATGACGCGCGCGCGCGCGCGCG 840
 QY 841 AAGAACTGCGCGCTGATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
 DB 841 AAGAACTGCGCGCTGATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
 QY 901 GACATGCGCATGAGCTCAACGAGATGATCCGCGAGTGGGATCTGCGAAGTACCGCT 960
 DB 901 GATATGCGCTTCAACTCAACGAGATGATCCGCGAGTGGGATCTGCGAAGTACCGCT 960
 QY 961 TTGCGGCTATGAGCACTCTTTCGCGAGTGTGCGCACTACTACGATGCGAGCGCG 1020
 DB 961 TTGCGGCTATGAGCACTCTTTCGCGAGTGTGCGCACTACTACGATGCGAGCGCG 1020
 QY 1021 GAGCTGCGCGAGGACATCGACATGAGCTGAGACCGCGCGCGCGCGCGCGCGCG 1080
 DB 1021 GAGCTGCGCGAGGACATCGACATGAGCTGAGACCGCGCGCGCGCGCGCGCGCG 1080

QY 1081 ATGCTGATGCTGCGGAGGAGCATGCCCGGTGCGCGCGCGCGCGCGCGCGCGCG 1140
 DB 1081 ATGCTGATGCTGCGGAGGAGCATGCCCGGTGCGCGCGCGCGCGCGCGCGCGCG 1140
 QY 1141 ATGCTGCGGAGAGAGCGGTGCGCGAGACATCAACGCGCTTCGCTGCGTGGCA 1200
 DB 1141 ATGCTGCGGAGAGAGCGGTGCGCGAGACATCAACGCGCTTCGCTGCGTGGCA 1200
 QY 1201 ATCATCCGCAAC 1212
 DB 1201 ATCATCCGCAAC 1212
 RESULT 7
 US-10-251-078-21
 ; Sequence 21, Application US/10251078
 ; Publication No. US20030119084A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Roche Diagnostics GmbH
 ; APPLICANT: Shao, Zhixin
 ; APPLICANT: Schmuck, Rainer
 ; APPLICANT: Kratzsch, Peter
 ; APPLICANT: Kenkies, Janet
 ; APPLICANT: Weisner, Harald
 ; TITLE OF INVENTION: Variants of an Erwinia-type creatinase
 ; FILE REFERENCE: 20981 EP
 ; CURRENT APPLICATION NUMBER: US/10/251,078
 ; CURRENT FILING DATE: 2002-09-20
 ; NUMBER OF SEQ ID NOS: 32
 ; SOFTWARE: Patentin version 3.1
 ; SEQ ID NO 21
 ; LENGTH: 1212
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: variant Ctsd7
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(1212)
 US-10-251-078-21
 Query Match 80.7%; Score 978.4; DB 15; Length 1212;
 Best Local Similarity 88.0%; Pred. No. 6,7e-257;
 Matches 1066; Conservative 0; Mismatches 146; Indels 0; Gaps 0;
 QY 1 ATGACTGACGACATGTTGCACTGATGAAATGGCAACAAGCGGAGAAATTAATGCGG 60
 DB 1 ATGACTGACGACATGTTGCACTGATGAAATGGCAACAAGCGGAGAAATTAATGCGG 60
 QY 61 TTTTCGGAATGCGAGATGACCGCGCGCAAAAAGAGCTTCCGCGTGTGATGCGCAAGAAC 120
 DB 61 TTTTCGGAATGCGAGATGACCGCGCGCAAAAAGAGCTTCCGCGTGTGATGCGCAAGAAC 120
 QY 121 AATGTGATGCGGCGCTGTTCACTCTTATCACTGATCAACTACTATTCGCGCTG 180
 DB 121 AATGTGATGCGGCGCTGTTCACTCTTATCACTGATCAACTACTATTCGCGCTG 180
 QY 181 TACTGCTATTTGCGAGCGCAAGTACGCGATGATGACCAACAACGCGACGACGATT 240
 DB 181 TACTGCTATTTGCGAGCGCAAGTACGCGATGATGACCAACAACGCGACGACGATT 240
 QY 241 TCGGCGGCGATGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
 DB 241 TCGGCGGCGATGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
 QY 301 ACCGACTGCGCGCGCGCGCAATTTCTATCGCGCGCGCGCGCGCGCGCGCGCGCG 360
 DB 301 ACCGACTGCGCGCGCGCGCAATTTCTATCGCGCGCGCGCGCGCGCGCGCGCGCG 360
 QY 361 CGGATGCGGATGCGAGTTGCAACGCGCACTGCGACTTCGCGCGCGCGCGCGCGCG 420
 DB 361 CGGATGCGGATGCGAGTTGCAACGCGCACTGCGACTTCGCGCGCGCGCGCGCGCG 420

421 CTACCGGCGCTGAGTTCGTGACATCAAGCCCTCGATGTGATGCGCACATCAAG 480
421 CTGCGCGGCGTGAAGTTCGTGATATCGGTCAACCGTCAATGTGATGCGCACATCAAG 480
481 TGGCTCGAAGAGAGAGAGTGTATCGCGAAGCGCGCGGTGTGTGACGTGCGCGCGG 540
481 TGGCTCGAAGAGAGAGAGTGTATCGCGAAGCGCGCGGTGTGTGACGTGCGCGCGG 540
541 GCGTGGCGCGCTGCGCATCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
541 GCGTGGCGCGCTGCGCATCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
601 AATGCGATGATCCGCGAGATCCGCAATGCTTCCCTTGTGTGAGCTGATGACACTG 660
601 AATGCGATGATCCGCGAGATCCGCAATGCTTCCCTTGTGTGAGCTGATGACACTG 660
661 ACCGTGTTCAAGTCGCGCATCAACCGACGCGCGCGCGCGCGCGCGCGCGCGCG 720
661 ACCGTGTTCAAGTCGCGCATCAACCGACGCGCGCGCGCGCGCGCGCGCGCGCG 720
721 GTGCAATCCGCGCGCATCTCTTTCGTGACACCTTCCCGATGATCTTCCGCTACTAC 780
721 GTGCAATCCGCGCGCATCTCTTTCGTGACACCTTCCCGATGATCTTCCGCTACTAC 780
781 GCGCTGAGCGGACGCTGTTCTGCGACATGTGATGATGCGCGCGCGCGCGCGCG 840
781 GCGCTGAGCGGACGCTGTTCTGCGACATGTGATGATGCGCGCGCGCGCGCGCG 840
841 AAGAACGTGCGCGTGCATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
841 AAGAACGTGCGCGTGCATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
901 GACATCGCATGATGAGCTCAACGAGATGTACCGCGAGTGGAGACTGTGAAATAC 960
901 GATATCGCATGATGAGCTCAACGAGATGTACCGCGAGTGGAGACTGTGAAATAC 960
961 TTCGCTATGAGCACTCTTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1020
961 TTCGCTATGAGCACTCTTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1020
1021 GAGCTGCGGAGAGACATGACACCGAGCTGAAGCCCGCGCGCGCGCGCGCGCG 1080
1021 GAGCTGCGGAGAGACATGACACCGAGCTGAAGCCCGCGCGCGCGCGCGCGCG 1080
1081 ATGTGATGATGCTCCCGAGAGGCGATGCGCGCGCGCGCGCGCGCGCGCGCG 1140
1081 ATGTGATGATGCTCCCGAGAGGCGATGCGCGCGCGCGCGCGCGCGCGCGCG 1140
1141 ATGCTGCGGAGAGAGAGCGTGCAGAGACATCAACCGCGCTGCGCGCGCGAG 1200
1141 ATGCTGCGGAGAGAGAGCGTGCAGAGACATCAACCGCGCTGCGCGCGCGAG 1200
1201 ATCATCCGCGAAC 1212
1201 ATCATCCGCGAAC 1212

RESULT 8
US-10-251-078-19
; Sequence 19, Application US/10251078
; Publication No. US20030119084A1
GENERAL INFORMATION:
; APPLICANT: Roche Diagnostics GmbH
; APPLICANT: Shao, Zhixin
; APPLICANT: Schmeck, Rainer
; APPLICANT: Kretsch, Peter
; APPLICANT: Krenke, Janet
; APPLICANT: Weisner, Harald
TITLE OF INVENTION: Variants of an Erwinia-type creatinase
FILE REFERENCE: 20981 EP
CURRENT APPLICATION NUMBER: US/10/251,078
CURRENT FILING DATE: 2002-09-20
NUMBER OF SEQ ID NOS: 32

SOFTWARE: PatentIn version 3.1
SEQ ID NO 19
LENGTH: 1212
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: variant CTs82
NAME/KEY: CDS
LOCATION: (1)..(1212)
US-10-251-078-19
Query Match 80.6%; Score 976.8; DB 15; Length 1212;
Best Local Similarity 87.9%; Pred. No. 1.8e-256;
Matches 1065; Conservative 0; Mismatches 147; Indels 0; Gaps 0;
1 ATGATGACGACATGTTGACGATGATGAATGACCAACGCGGAGAAATATTCGCGG 60
1 ATGATGACGACATGTTGACGATGATGAATGACCAACGCGGAGAAATATTCGCGG 60
61 TTTTGGATGCGCGAGATGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
61 TTTTGGATGCGCGAGATGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
121 AATGTCGATGCGCGCGCTGTTCACTCTTATCACTGATCACTAATTCGCGCTG 180
121 GACGTGACGCGCGCTGTTCACTCTTATCACTGATCACTAATTCGCGCTG 180
181 TACTGCTATTTTCGACGCGAGATGCGGATGATGATGACGACGACGACGACGAT 240
181 TACTGCTATTTTCGACGCGAGATGCGGATGATGATGACGACGACGACGACGAT 240
241 TCGCGCGCGCATGAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
241 TCGCGCGCGCATGAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
301 ACCGATGCGCGCGCGCGCAATTTCTATGCGCGCGCGCGCGCGCGCGCGCG 360
301 ACCGATGCGCGCGCGCGCAATTTCTATGCGCGCGCGCGCGCGCGCGCGCG 360
361 GCGATGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
361 GCGATGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
421 CTACCGGCGCTGAGTTCGTGACATCAAGCCCTCGATGTGATGCGCGCGCG 480
421 CTACCGGCGCTGAGTTCGTGACATCAAGCCCTCGATGTGATGCGCGCGCG 480
481 TCGCTCGAAGAGAGAGAGTGTATCGCGAAGCGCGCGCGCGCGCGCGCGCG 540
481 TCGCTCGAAGAGAGAGAGTGTATCGCGAAGCGCGCGCGCGCGCGCGCGCG 540
541 GCGTGGCGCGCTGCGCATCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
541 GCGTGGCGCGCTGCGCATCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
601 AATGCGATGATCCGCGAGATCCGCAATGCTTCCCTTGTGTGAGCTGATGACACT 660
601 AATGCGATGATCCGCGAGATCCGCAATGCTTCCCTTGTGTGAGCTGATGACACT 660
661 ACCGTGTTCAAGTCGCGCATCAACCGACGCGCGCGCGCGCGCGCGCGCGCG 720
661 ACCGTGTTCAAGTCGCGCATCAACCGACGCGCGCGCGCGCGCGCGCGCGCG 720
721 GTGCAATCCGCGCGCATCTCTTTCGTGACACCTTCCCGATGATCTTCCGCTACTAC 780
721 GTGCAATCCGCGCGCATCTCTTTCGTGACACCTTCCCGATGATCTTCCGCTACTAC 780
781 GCGCTGAGCGGACGCTGTTCTGCGACATGTGATGATGCGCGCGCGCGCGCG 840
781 GCGCTGAGCGGACGCTGTTCTGCGACATGTGATGATGCGCGCGCGCGCGCG 840
841 AAGAACGTGCGCGTGCATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900

```

Db      ||||| 900
841 MGAAGTGCCTGTCACCGCGGCTTCGAACTCATCAAGCGGTGGCTGCAAG
Qy      ||||| 960
901 GACATCGCATGAGCTCAACGAGATGTAACCGGAGTGGACCTGTGTAAGACCGCTTC
Db      ||||| 960
901 GATATCGCATGAGCTCAACGAGATGTAACCGGAGTGGAGTGTGTGAAGTACCGCTTC
Qy      ||||| 1020
961 TTGGGTATGCGCATCTCTTCCGCGTGTGTGCTGCTACTAGGTGTGGAGGCGGCGT
Db      ||||| 1020
961 TTGGGTATGCGCATCTCTTCCGCGTGTGTGCTGCTACTAGGTGTGGAGGCGGCGT
Qy      ||||| 1080
1021 GAGCTGCGGAGGACATGCAACCGAGCTGAAGCCGGAGTGGTCTCAATGAGCGG
Db      ||||| 1080
1021 GAACTGCGGAGGACATGCAATCGTGTGCAAGCCGGAGTGGTCTCAATGAGCGG
Qy      ||||| 1140
1081 ATGATGATCTGCGGAGGACATGCGGTGCGGCGGCTATGCGAGACGACATCTGT
Db      ||||| 1140
1081 ATGATGATCTGCGGAGGACATGCGGTGCGGCGGCTATGCGAGACGACATCTGT
Qy      ||||| 1200
1141 ATGCTGCGGAGGACGCTGCGGAGAACATCACTGCTTCCGTTCCGTGCGGAGACGAC
Db      ||||| 1200
1141 ATGCTGCGGAGGACGCTGCGGAGAACATCACTGCTTCCGTTCCGTGCGGAGACGAC
Qy      ||||| 1200
1201 ATCATCGGAC 1212
Db      ||||| 1201
1201 ATCATCGGAC 1212

```

RESULT 9

US-10-437-963-88003/C

Sequence 88003, Application US/10437963

Publication No. US20040123343A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

APPLICANT: Wu, Wei

APPLICANT: Boukharov, Andrey A.

APPLICANT: Barbazuk, Brad

APPLICANT: Li, Ping

TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

FILE REFERENCE: 38-21(53221)B

CURRENT APPLICATION NUMBER: US/10/437,963

CURRENT FILING DATE: 2003-05-14

NUMBER OF SEQ ID NOS: 204966

SEQ ID NO 88003

LENGTH: 2733

TYPE: DNA

ORGANISM: Oryza sativa

FEATURE:

OTHER INFORMATION: Clone ID: PAT_MRT4530_86897C.1

US-10-437-963-88003

Query Match 6.1%; Score 73.8; DB 19; Length 2733;
 Best Local Similarity 44.3%; Pred. No. 6.9e-10;
 Matches 398; Conservative 0; Mismatches 492; Indels 9; Gaps 2;

```

Qy      ||||| 252
252 CGACGGCGGCGCCCTGCGCGCGGCACTTCCGCGCAACATCACTACACGACGCGGCG 311
Db      ||||| 2506
2506 CGACGGCGGCGGTAGCGCGCGGTGGGACGAGCGCTTCCACTTACTGCGCCCACTTCCG 2447
Qy      ||||| 312
312 CGCGCAATTTCTATCGCGCGCTGCGCACTGACCAAGCGGCGCAAGCGCATTCGGCAT 371
Db      ||||| 2446
2446 CGAGAACGTGCTTTCTCGGTCAAGGTGCGGTGTGCTGCTGCGTCAAGTCAAGTATCGGCG 2387
Qy      ||||| 372
372 CGAGTTGCAACAGCTCAATCTGCACTTCCGCGCGCAAGCTGAGAGAGCCCTACCGGCGT 431
Db      ||||| 2386
2386 CGCTACTCTCCCGTACGAGGACCTCTGTCGCGCGAGGCGCTGAGACGCAAGCTGACAT 2327
Qy      ||||| 432
432 CGAGTTGTCGACATACGACGCGCTGATGTGATGCGACCATCAAGTCTGCAAG 491

```

```

Db      ||||| 2267
2236 CCTGGGAGACGACAAAGAAAGCTCCCCACGCTCCGACATCCAGCTCCGCTGCAATT 2267
Qy      ||||| 551
492 GCAAGACTGATCCGCGAAGCGCCCGCTGTGTGACGTTGGCGGCGGCGCTGGCGGC 551
Db      ||||| 2207
2266 CAAGACGCTGGCGCGCGACCGCGCGGAGAGTGTGGGCGCGCGCGCTGCGAGCGGC 2207
Qy      ||||| 611
552 TGGCATCAAGCGCGCGTGGCGGAGCATGAATGGCGATTCGCGCACCAATATGCAATGAT 611
Db      ||||| 2147
2206 GTACCCCGCGTGGCGTGCATCTTACTTAAGACGACGCGCGGTGACAGGTCACTCTGT 2147
Qy      ||||| 671
612 CGCGAGATTCGCAATGTTTCCCTTGTGAGCTGATGACACATCGAATCTGTGTTCA 671
Db      ||||| 2087
2146 CCAGAGACCGGACCGCGCGGACAGGTTCGCGCGAGATCCGCTGCGCGGCGCGCGCA 2087
Qy      ||||| 731
672 GTGGGCAATCAACACGCGCGGCAATTCGCTACACACCGCATTCGTGCAATTCGG 731
Db      ||||| 2033
2086 CTACGAGAGGCGGCTGTGGAGAGATGTTTCACCCATCAGCAACG-----CCAA 2033
Qy      ||||| 791
732 CGACATCTTTCGCTCAACACCTTCGCGATGCTTTCGCTACTACACCGCGCTGAGCG 791
Db      ||||| 1973
2032 GCACTTCATATACATCACCGGCTGTGTGCTGCTTTCACCAATACAGCTCATTCGCGAC 1973
Qy      ||||| 851
792 CACGCTGTTCTGCGACATGTGATGACGCGACCTCGACATCTGGAGAAAGACGTGAC 851
Db      ||||| 1913
1972 GTCCGCGAGCGCGCGCGCGGAGCGCACCATCGGCGAGCTCTCAAGCGCAAGCGCAG 1913
Qy      ||||| 911
852 GTGTCATCGCGCGGCTGTGAGCTGATCAAGCCGCGCGCGCTGCAAGACATCGCAT 911
Db      ||||| 1853
1912 CGAGGCGCTGCGCGTGTCTGATGCTGTGTGAAACACATCTTCTTCATTCAGGCGCTCA 1853
Qy      ||||| 971
912 CGAGCTCAACGAGATGTAACCGGAGTGGACCTGTGAAATGACCGCTCTTTCGCTATG 971
Db      ||||| 1793
1852 GCGCATTCGCAATCAAGCTCAAGCTGTGCGCGACGACGACGACGACGACGCTGCTACT 1793
Qy      ||||| 1031
972 CCACTCTTTCGCGCTGTGTGCACTACTAGTGTGCGAGGCGGCGGTGAGACTGCGCA 1031
Db      ||||| 1733
1792 CGAGGACTCCAGCGTGCATGCTGCTGTGCTGCGCGCGGCAAGCGCGCGCGCGCAG 1733
Qy      ||||| 1088
1032 GACATTCACACCGAGCTGAAGCCCGGCAATGTGTCTTCATGAGCCGATG---GTAT 1088
Db      ||||| 1673
1732 CAGCTTCATCAATGGGCAAGAGTCTCACTTTCGCGACCCACGACGCAAGAACCGTCA 1673
Qy      ||||| 1147
1089 GGTGCGGAGGCGGAGCGCGGTGCGGCGCTATGCGGACGACGACATCTGATGCTG 1147
Db      ||||| 1614
1672 CGTGACCAACGACATGCGCGCGCGGACCGGCGCGGCGGCTCGGACATCCGCTGATCG 1614

```

RESULT 10

US-10-425-115-109294/C

Sequence 109294, Application US/10425115

Publication No. US20040214272A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

FILE REFERENCE: 38-21(53222)B

CURRENT APPLICATION NUMBER: US/10/425,115

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 369326

SEQ ID NO 109294

LENGTH: 1262

TYPE: DNA

ORGANISM: Zea mays

FEATURE:

OTHER INFORMATION: Clone ID: MRT4577_31174C.1

US-10-425-115-109294

Query Match 5.9%; Score 71.8; DB 20; Length 1262;
 Best Local Similarity 43.9%; Pred. No. 2.1e-09;
 Matches 307; Conservative 0; Mismatches 392; Indels 0; Gaps 0;

```

Oy 205 GGCATGATCATGCAACCAACACGCGACGATTCGCGCGCATGACGCGCGCAG 264
    |||
Db 1107 GGCAGCGGCGCTCCGCTCCAAAGACGATGTCACCTTACAAAGTCCAAACGCGCATCGCC 1048
Oy 265 CCGTGGCGCGGAGCTTCGCGCGCAACATCACTTACACGCGCTGCGCGCGCAATTC 324
    |||
Db 1047 ACGATGCGCTTCTTACTTCCCGACGACCTTCTTCAACGCGCGCTGCGCGCATTC 988
Oy 325 TATCGCGCGGTGGCGACGTCACACGCGCGCGCAACGCGCATTCGAGTTCCGAC 384
    |||
Db 987 GCGCTGCGCTTGAAGTCCGCAACGCGCAAGTGGGAGCTGCGCGCGCGCGCGCGCAG 928
Oy 385 GTCAATCTGCACTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 444
    |||
Db 927 GCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 868
Oy 445 ATCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 504
    |||
Db 867 GTGCTGCTGCGGAAACGAGGTGCTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 808
Oy 505 CGCGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 564
    |||
Db 807 GTGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 748
Oy 565 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 624
    |||
Db 747 AAGATGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 688
Oy 625 AATGCTTCCCTTCTGTCGATGATGACACCTGGAACCTGTCAGTCGCGCGCATCA 684
    |||
Db 687 GCGATGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 628
Oy 685 ACCGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 744
    |||
Db 627 TACCCCTTCTTCTGTCGATGATGACACCTGGAACCTGTCAGTCGCGCGCATCA 568
Oy 745 CTCAACACCTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 804
    |||
Db 567 CAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 508
Oy 805 GACCATGTCGATGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 864
    |||
Db 507 GACCGATGTCGATGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 448
Oy 865 GCGCTGAGCTGATCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 903
    |||
Db 447 GACGTCGTGTCGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 409

```

RESULT 11
US-10-437-963-38426
Sequence 38426, Application US/10437963
Publication No. US20040123343A1

GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Mu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
CURRENT APPLICATION NUMBER: US/10/437,963
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 38426
LENGTH: 1483
TYPE: DNA
ORGANISM: Oryza sativa

FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_42062C.1
US-10-437-963-38426

Query Match 5.8%; Score 69.8; DB 19; Length 1483;
Best Local Similarity 4.9%; Pred. No. 7.6e-09;
Matches 266; Conservative 0; Mismatches 327; Indels 0; Gaps 0;

```

Oy 587 GATGCGCAACCAACGATGATCCGCGAATGCGCGCGCGCGCGCGCGCGCGCGCGCG 646
    |||
Db 309 GAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 368
Oy 647 TGATGACACCTGACCTGCTTCAATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 706
    |||
Db 369 TCGCGGACCTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 428
Oy 707 TCACCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 766
    |||
Db 429 ATGACCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 488
Oy 767 TCGGCTATCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 826
    |||
Db 489 GCGCGGCTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 548
Oy 827 TCGCATCTGGGAGAAAGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 886
    |||
Db 549 ACATGCGCATCACTGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 608
Oy 887 GCGCGGCTGCAAGGACATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 946
    |||
Db 609 TCTGCTACGTCAACGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 668
Oy 947 TGAAGTACCGCTCTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1006
    |||
Db 669 TCATCTAGTGAATGACATGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 728
Oy 1007 GCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1066
    |||
Db 729 CCAACCGGATGATGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 788
Oy 1067 TCTCCATGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1126
    |||
Db 789 TCTTCCGCGGCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 848
Oy 1127 AGCAGACATCTGATGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1179
    |||
Db 849 TGAACGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 901

```

RESULT 12
US-10-411-910A-222
Sequence 222, Application US/10411910A
Publication No. US20040209256A1

GENERAL INFORMATION:
APPLICANT: Dillon, Harrison F.
TITLE OF INVENTION: Methods and Compositions for Evolving Hydrogenase Genes
FILE REFERENCE: H2041203-P
CURRENT APPLICATION NUMBER: US/10/411,910A
NUMBER OF SEQ ID NOS: 343
SOFTWARE: PatentIn version 3.2
SEQ ID NO 222
LENGTH: 1212
TYPE: DNA
ORGANISM: Scenedesmus obliquus
US-10-411-910A-222

Query Match 5.7%; Score 69.6; DB 20; Length 1212;
Best Local Similarity 4.2%; Pred. No. 8.3e-09;
Matches 363; Conservative 0; Mismatches 489; Indels 0; Gaps 0;
Oy 183 CTGCTATTTGAGCAAGTACGCGCATGTCATGACACAAACGCGACGATTC 242
 |||


```

Db      294 CAGCTGTGCCCCGCTGGGTGGCCATGGTGGAGGAAGCAACCCGAGCTATCCCTTA 353
Qy      243 GGGCGGCATGACGGGGCCAGCCCTGGGGCCGAGCTTGGCGAACAATCACTCAAC 302
Db      354 CCGAGACAGCTGGAGAGCCCGCAGATGATGCTGGGGCCGCTGATCAAGAACTACTTCCG 413
Qy      303 CGACTGGGGCCGCAATTTCTATGCGCCGTGGCCCAAGCTGACCAACGGGGCCGAAGCG 362
Db      414 CGCCGAGCCCGGGCCGAGCCCGAGGACATCTGCAAGTGAAGCTGATGCTGCTGG 473
Qy      363 CATCGGCATGAGTTTGAACACGTCATCTGACTTTCGCGCCGACCTGAGAGAACCTT 422
Db      474 CAAGAGCGGGAGCGGAGCCCGAGCCCGAGCAACCAACCGCGCGCGCGCGCA 533
Qy      423 ACCGGGCGTGAAGTTTGTGACATGAGCCAGCCCTTGAATGATGATGCAACCTCAAGTC 482
Db      534 CGTGAACAAGTATACCAACGCGCCGAGCTGGGCAAGATCTTGTGAGAGCGCGCACTCA 593
Qy      483 GCTCGAAGAGCAAGACTGATTCGCGAAGGCGCCGCGTGTGATGAGTGGGGCGCGG 542
Db      594 GCTTGAACAGCTGACAGAGAGCCCTTTCACAAACCCGTGGGCGAGAGGCAAGCGCGCG 653
Qy      543 CTGCGGCGCTGCATCAAGCGCGCGCTGCGCGAGCATGAAGTGGCGATCGCAACCA 602
Db      654 CCGCTGTTCCGGACCAACCGGGGGCGGTGATGAGAGCGCGCGCTGCGACCGTGAAGAGT 713
Qy      603 TGGCATGATCCCGAGATGCGCAATGCTTCCCTTGTGAGAGCTATGAGCACTTGAAC 662
Db      714 GGTGACCGCGAGGCGCTGGGCGCCCGAGCGCAAGCTGACCAACGAGACCGCTGAGC 773
Qy      663 CTGGTTCCAGTGGGGCATCAACCGAGCGCGCAATTCGGGTCAACCAACCGCATCGT 722
Db      774 CCCCCCGCCAGCGGCGAGCCCGCGCCCGCCAGCCCGCCAGCCCGCCAGCCCGCC 833
Qy      723 GCATCCCGGCAATCTTTCCTCAACACTTCCCGATGATCTTGGCTACTTCAACCGC 782
Db      834 CCGTGAAGGCGCAAGCCGAGAGCGGATCAACCTGAACATCGCGTGAAGCGGCTGGG 893
Qy      783 GCTGAGAGCGCAAGCTTCTTGTGACCAATGTCATGAGCGCAGCTTGAACATCTGGAGAA 842
Db      894 CAACCCCAAGAACTGATCAAGCAAGCTGCGCCCGCGAGAGCAAGTACGACTTCAACCGA 953
Qy      843 GAACGTCGCTGATCAAGCGCGCGGCTGAGCTGATCAAGCGGGCGCGCGTCAAGGA 902
Db      954 GGTGATGCGCTTCCCGCGGCTGATCGAGCGGCGCGCGCGCGCGCGCGCAACAGCA 1013
Qy      903 CATCGGCATGAGCTCAAGAGATGTAACCGCGAGTGGGACTGCTGAATACCGCTCTT 962
Db      1014 GATCTCGAAGAAAGCGCGAGCGCGCATGTAACGACTGGAAGAGCGCGCGTGAATCGCGCG 1073
Qy      963 CGGCTATGCGCACTCTTGGGGGTGTGTCCTACTACGCTGCGGAGCGCGGCTGGA 1022
Db      1074 CACCGAAGAACCCCTGATGCGGCGCTTGAAGAGATCTTGGGAGAGCCCAACCGCGCA 1133
Qy      1023 GCTGCGCGAGGA 1034
Db      1134 CAGGCGCCACGA 1145

```

```

RESULT 13
US-10-437-963-57011/c
; Sequence 57011, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Bardazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

```

```

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 57011
; LENGTH: 1169
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_58865C.1
US-10-437-963-57011

```

```

Query Match      5 7%; Score 68.6; DB 19; Length 1169;
Best Local Similarity 45.3%; Pred. No. 1,6e-08;
Matches 248; Conservative 0; Mismatches 299; Indels 0; Gaps 0;

```

```

Qy      512 GCGGCGCGGCGCTGCGCGCTGCGCATCAAGCGCGCGTCCGAGCATGAGTGGCATC 591
Db      918 GCGGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 859
Qy      592 GCCACCAACCATGCGATGATCCGAGATCGCCAAATGCTTCCCTTGTGAGCTGATG 651
Db      858 AACCTGCTCCCGCGCGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 799
Qy      652 GACACCTGAGCCTGCTTCAAGTGGGCAATCAACCGAGCGCGCGCAATTCGGTCAAC 711
Db      798 GTGGGCTTGGGGGCGAGCAAGCTGAGAGCAACCGCGCGCGCGCGCGCGCGCGCG 739
Qy      712 AACCCATGTCATTCGCGCGCGCATCTTTCCTTCAACACTTCCCGATGATCTTCCG 771
Db      738 GTGACGTGTGGTGGCGCAAGCGCGCGTGGCTGCTGCTGCGCATGAGCGCTTACGCG 679
Qy      772 TACTAACCGCGCTGAGCGCAAGCTGTTCTGCGACATGTGATGAGCGCGCGCTGAC 831
Db      678 CTGACGCGCGTGAAGTCAAGTCAAGCACTTCAACGAGCGCGCGCGCGCGCGCG 619
Qy      832 ATCTGGGAAGAAAGTGGCGGTGATGCGCGCGCGCGCTGAGCTGATCAAGCGCGCGCG 891
Db      618 AGTTCGTGAGTGAATGATGCGCGCGCTTCTTCAACAGCTTCAAGCGCGCGCGCGCG 559
Qy      892 CGCTCAAGAGCATCGCATGAGCTCAACAGATGTAACCGAGTGGGAGCTGCTGAG 951
Db      558 ACCACTTCATGCGCGCGCTTGAAGAGCGCGGTGAGAGCGCGTACTACAGCGCGCTG 499
Qy      952 TACCGCTCTTGGGTATGAGCGCATCTTTCGCGGTGCTGTGCACTACTAGCGTGG 1011
Db      498 CGGCGCTACGCGCGCGGTGATGACTCTGCTCAACTTCCAGTTCTACGCGCTACG 439
Qy      1012 GCGCGCGTGAAGCTGCGGAGAGCAATGCAACCGAGCTGAAGCGCGCGCATGCTGCTC 1071
Db      438 ACCGAGTGGCGAGCTAGTATGTTCTTACAGAGAGCGCGCGCATACCGCGCGCG 379
Qy      1072 ATGAGCG 1078
Db      378 AAGTGC 372

```

```

RESULT 14
US-10-156-761-4932
; Sequence 4932, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761

```

```

; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 4932
; LENGTH: 834
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(834)
; US-10-156-761-4932

```

```

Query Match      5.6%; Score 68; DB 15; Length 834;
Best Local Similarity 46.6%; Pred. No. 2.1e-08;
Matches 332; Conservative 0; Mismatches 365; Indels 15; Gaps 3;

```

```

QY 475 ATCAAGTCGCTCGAAGAGCAGAAAGCTGATCCGCGAAGCGCCCGCGTGTGTGAGTGGC 534
DB 10 ATCAAGTCGCTCGAAGAGCAGAAAGCTGATCCGCGAAGCGCCCGCGTGTGTGAGTGGC 69
QY 535 GGGCGGCGCTGCGCGGCTGCAATCAAGCGCGGCTGCGCGAGCATGAGTGGCATGCGC 594
DB 70 CATGCGGCGCACTGATGAGCGCGGCTGCGCGGCGCGCACTGAGAGCATGAGTGGC 129
QY 595 ACCACCAATGCAATGATCCGCGAGATGCGCAATTCCTTCCCTTCTGAGCTGATGAC 654
DB 130 GCGCGGAGAGTGTCCGCGAGCAGCGTGGCGAAGTTCGAACTTCTCGGCTTCCGCGCTTC 189
QY 655 ACCTGAGCTGTTTCAAGTGGGATCAACACGAGCGGCGGCAATCCGCTGACCAAC 714
DB 190 CCGCGCAATCTGCACTGCTGCTCAACGAGTGTGTGTGATGATCCGCGAGCAAG 249
QY 715 CGCATGTCATTCGCGCGCAATCTTTCGCTCAACCTTCCGAGTATCTTTCGCTAC 774
DB 250 ACCGTCCTGAAGGACGCGCGAGTCTCATCTCGATGACGCGGCGGCGATTCAGCGGCTGG 309
QY 775 TACACGCGCTGAGAGCGCAAGCTTCTGAGCACTGATGATACCGCAACC---TCGAC 831
DB 310 CAGCGGAGCGCGCTTACACCGCTTCTGAGGTTCCGCGCACTCTCGAGCTGTCGAG 369
QY 832 ATCTGAGAGAGAAAGTGGCGGCTGATCCGCGGCGGCTGAGCTGATCAAGCGCGGCGG 891
DB 370 CTCTCCCGGAGTGGCGGAGAGTGTGATGAGCGGCGCATGCGCGAGTGAAGAGGCGAAC 429
QY 892 CGCTGAGAGAGATTCGCAATCGAGTCAACGAGATGATCGCGAGTGGGACTGCTGAAG 951
DB 430 CGCTGAGTGAAGTGTCTCCGCTGCAATCGAGTCAATCGCGCGAGCGCGAGCGGCGG 489
QY 952 TACCGCTCTTGGGCTATGAGCCACTCTTCGCGGCTGCTGCTGCACTAACGCTGCGAG 1011
DB 490 GCGCGGAGAGTACGCGGAGTCTGCGAGACTTACGCGG-----GCCACGCGCATCGGCA 543
QY 1012 GCGCGGAGTGAAGTGGCGGAGCAATCGACACCGAGCTGAAGCGCGGAGTGTGTCTCC 1071
DB 544 ATGCAATGAGAGCCGCACTGCTGAACCTTGTGAGACCGCGCGCGCAAGAGGTTCGAG 603
QY 1072 ATGAGAGCGATGATGATGCTGCGCGAGGAGATGCGCGGCTGATGCTGAGCGAGAC 1131
DB 604 CTGCTGCGCGGCTTCTGCTGCGAGTCAAGCGGAGTGTGTCTCTCGGCACTGAGAGAC 663
QY 1132 GACATCTGATCTGCGGAGAGAGCGGTGCGAGAACATCAACCGGCTTTCCTG 1183
DB 664 GA-----GCTGCTGAGAGAGAGCTGAGACCGTCACTCAACCGAGCGGCAAGT 709

```

```

RESULT 15
US-10-363-345A-31407/c
; Sequence 31407, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Alexander Olek
; APPLICANT: Christian Pienbrock
; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; FILE OF INVENTION: cytosines in genomic DNA in the sequence context of 5'-CpG-3
; PILE REFERENCE: 501/1227
; CURRENT APPLICATION NUMBER: US/10/363,345A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 31407
; LENGTH: 985
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CpG-Island No: 31407
; US-10-363-345A-31407

```

```

Query Match      5.6%; Score 68; DB 20; Length 985;
Best Local Similarity 43.0%; Pred. No. 2.2e-08;
Matches 385; Conservative 0; Mismatches 510; Indels 1; Gaps 1;

```

```

QY 152 ACTGATCACTAATTCGCGCTGCTGATCTGATTTTCGAGCGCAAGTACGCGATG 211
DB 937 ACATGACATCAACATCTACGTGATGATCATCTACGTGACCTGACATCTACGTGACATG 878
QY 212 TCAATCAACCAACAGCGGAGCATTTTCGCGGATTCAGCGGCGGCGGCGGCGGCGG 271
DB 877 ACATCTACATCTACGTGACATGACATCTACGTGACATCTACGTGACATCTACGTGAC 818
QY 272 GCGCGAGCTTGGCGGCAATCACTACCGAGCTGAGCGGCGGCGGCAATTTTCATGCG 331
DB 817 ACATGACATCAACATCTACGTGATGATGATGATGATGATGATGATGATGATGATGATG 758
QY 332 CCGTGGC-CAGCTGACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 390
DB 757 ACATCTACATCAACATCTACGTGATGATGATGATGATGATGATGATGATGATGATGATG 698
QY 391 CTGATCTTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 450
DB 697 TACATGACATCAACATCTACGTGATGATGATGATGATGATGATGATGATGATGATGATG 638
QY 451 CAGCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 510
DB 637 CCGCTTACATCTACGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 578
QY 511 GCGCGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 570
DB 577 AACGTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 518
QY 571 CCGAGAGATGAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 630
DB 517 GAGCTTACATCTACGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 458
QY 631 TTCCCTTGTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 690
DB 457 TACATCTACGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 398
QY 691 GCGCGGCAATTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 750
DB 397 AACGTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 338
QY 751 ACCCTCCGAGATCTTGGCTACTACACCGCGCTGAGAGCGGCGGCGGCGGCGGCGGCGG 810
DB 337 GACATGAGAGTGTACATCTACGTGATGATGATGATGATGATGATGATGATGATGATGATG 278
QY 811 GTGATGAGCGGAGCGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 870
DB 277 TACGTCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 218
QY 871 GAGCTGATCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 930
DB 217 GACATCAACATCTACGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 158

```

Oy 931 CGGAGTGGGACCTGCTGAAGTACCGCTCCTTGGCTATGGCCACTCCCTTGGCGGCTG 990
Db 157 TACGTGACGCTCTACATCTACGTGACGTGACGCTCTACATCAACATCAACGTGACGTC 98
Oy 991 TGGCACTACTACGCTGCGAGGCCGCGCTGAGACTGCGAGGACATCGACACCGA 1046
Db 97 TACATCAACATCTACATCAACGTGACGTGACGTCTACATCAACGTGACGCTTA 42

Search completed: July 8, 2005, 01:32:51
Job time : 839 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 7, 2005, 23:30:26 ; Search time 81 Seconds
(without alignments)
1929.027 Million cell updates/sec

Title: US-10-807-228a-1

Perfect score: 2190

Sequence: 1 MTDDMLHVMKHNKGEKDYSP.....DGAENITGFPEGHNIIRN 404

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2190	100.0	404	2	AAW11861 Creatinin
2	2190	100.0	404	2	AAW22893 Creatine
3	2190	100.0	404	2	AAW61905 Scabie cr
4	2190	100.0	404	3	AAW09976 Alcalligen
5	2190	100.0	404	4	AAU08727 Creatine
6	2190	100.0	404	5	AAW51471 Alcalligen
7	2190	100.0	404	5	ABW05660 Creatine
8	2149	98.1	404	2	AAW94463 Creatine
9	2149	98.1	404	3	AAW12775 Alcalligen
10	2078	94.9	404	6	ABR43478 Mutant Er
11	2078	94.9	404	6	ABR43467 Erwinia s
12	2068	94.4	404	6	ABR43480 Mutant Er
13	2068	94.4	404	6	ABR43474 Mutant Er
14	2066	94.3	404	6	ABR43477 Mutant Er
15	2064	94.2	404	6	ABR43475 Mutant Er
16	2062	94.2	404	6	ABR43476 Mutant Er
17	2058	94.0	404	6	ABR43479 Mutant Er
18	1355.5	61.9	403	1	AAW60274 Sequence
19	1345.5	61.6	403	1	AAW7502 Wild type
20	1345.5	61.4	403	1	AAW80680 Creatine
21	1339.5	60.7	411	2	AAW80133 Arthrobac
22	1048.5	47.9	378	1	AAW70542 Sequence
23	270	12.3	353	7	ADW92484 B. lichen
24	255.5	11.7	353	2	AAW98041 Bacillus
25	248.5	11.3	351	4	AAW96572 Putative

26	248.5	11.3	351	8	ADW43202 Bacterial
27	246.5	11.3	351	8	ADW59244 Pyrococu
28	244	11.1	351	8	ADW47444 Thermoco
29	236.5	10.8	352	7	ADW86421 Enteroco
30	230.5	10.5	365	5	ABW48053 Listeria
31	229.5	10.5	354	4	AAW35052 Enteroco
32	229	10.5	361	4	AAW34642 E. coli c
33	223.5	10.2	353	5	ABW9480 Listeria
34	219	10.0	356	7	ADW9465 E. faeciu
35	219	10.0	359	6	ADW09220 Altioco
36	215.5	9.8	352	5	ABW54004 Lactoco
37	204.5	9.3	352	2	AAW98042 Lactoco
38	202	9.2	288	6	ADW09340 Altioco
39	202	9.2	353	6	ADW09342 Altioco
40	202	9.2	370	6	ADW09344 Altioco
41	200.5	9.2	370	5	ABW38924 Staphyloc
42	200.5	9.2	370	8	ADW06651 Staphyloc
43	200	9.1	350	4	AAW3791 Staphyloc
44	200	9.1	353	4	AAW36987 Staphyloc
45	200	9.1	365	6	ABW71039 Staphyloc

ALIGNMENTS

RESULT 1
ID AAW11861 standard; protein; 404 AA.
XX
AC AAW11861;
XX
DT 17-APR-1997 (first entry)
XX
DE Creatinine amidinohydrolyase.
XX
KW Thermal stability; creatinine amidinohydrolyase; creatine; sarcosine;
KM urea; blood.
XX
OS Alcalligenes faecalis.
XX
PN JP08308579-A.
XX
PD 26-NOV-1996.
XX
PF 16-MAY-1995; 95JP-00117283.
XX
PR 16-MAY-1995; 95JP-00117283.
XX
PA (TOYM) TOYOBO KK.
XX
DR WPI; 1997-059698/06.
XX
N-PSDB; AAW61367.
XX

Gene coding for creatinine amidinohydrolyase - used to quantify blood or urinary creatinine as a disease indicator.

Claim 1; Page 9-10; 12pp; Japanese.

This sequence represents a thermally stable creatinine amidinohydrolyase which has a low Km value for creatine. The creatinine amidinohydrolyase has the following physicochemical properties: (a) action: converts creatine and water into sarcosine and urea; (b) optimum temp.: 40-45 deg C; (c) optimum pH: 8.0-9.0; (d) thermal stability: about 50 or less (pH 7.5, 30 mins); (e) stable pH: a pH range of about 4-10; (f) Km value: about 15.2 mM; (g) mol. wt.: about 4300 (SDS-PAGE); and (h) isoelectric point: about 3.5. The creatinine amidinohydrolyase may be used in the quantification of blood or urinary creatinine and creatine as an indicator of various diseases

Query Match 100.0%; Score 2190; DB 2; Length 404;
Best Local Similarity 100.0%; Pred. No. 1.5e-214;

Matches 404; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTDDMLHVKMHNGEKDYSPFSDAEMTRONDVGMMAKNVDAALFTSYHCINYSGL 60
 DB 1 MTDDMLHVKMHNGEKDYSPFSDAEMTRONDVGMMAKNVDAALFTSYHCINYSGL 60

QY 61 YCYFGRKYGWVIDHNNATTISAGIDGQPMRRSFGDNITYTDWRDNFRAVQLTGAK 120
 DB 61 YCYFGRKYGWVIDHNNATTISAGIDGQPMRRSFGDNITYTDWRDNFRAVQLTGAK 120

QY 121 RIGIEFDHNLDPFRQLEALPGVEFVDISQPSMMRTIKSLEOKLIREGARVCVGA 180
 DB 121 RIGIEFDHNLDPFRQLEALPGVEFVDISQPSMMRTIKSLEOKLIREGARVCVGA 180

QY 181 ACAAIKAGVBEHEVAIATTNAMIIRIAKSPPEVLMOTWTWFOGINTDGANPVNRI 240
 DB 181 ACAAIKAGVBEHEVAIATTNAMIIRIAKSPPEVLMOTWTWFOGINTDGANPVNRI 240

QY 241 VQSGDILSLNTPMIFGYTALERTLPCDHVDASLDIWEKNAVHRGELIKPGARCK 300
 DB 241 VQSGDILSLNTPMIFGYTALERTLPCDHVDASLDIWEKNAVHRGELIKPGARCK 300

QY 301 DIAIELNEMRYEMDLKYRSFGYHSGFVLCYHGREAGVELREDIDTELKPGVVSMEP 360
 DB 301 DIAIELNEMRYEMDLKYRSFGYHSGFVLCYHGREAGVELREDIDTELKPGVVSMEP 360

QY 361 MWMLPEGMPGAGGYREHDILIVGEDGAENITGPPFGEHNIIIN 404
 DB 361 MWMLPEGMPGAGGYREHDILIVGEDGAENITGPPFGEHNIIIN 404

RESULT 2
 AAM22893 ID AAM22893 standard; protein; 404 AA.

AC AAM22893;
 DT 17-OCT-2003 (revised)
 DT 02-MAR-1998 (first entry)

DE Createine amidinohydroxylase enzyme.
 KM Createine amidinohydroxylase enzyme; sarcosine; urea; dye; Km.

OS Alcaligenes faecalis; - strain TE3581 (FERM P-14237).

FH Key Location/Qualifiers
 FT Misc-difference 145 /note= "encoded by GAC"

XX BEP90303-A1.
 XX 20-AUG-1997.
 XX 13-FEB-1997; 97EP-00102270.
 XX 13-FEB-1996; 96JP-00025435.
 XX (TOYM) TOYO BOSEKI KK.
 XX Sogabe A, Hattori T, Nishiya Y, Kawamura Y;
 XX WPI; 1997-404731/38.
 XX N-PSDB; AAT38807.
 XX Createine amidinohydroxylase enzyme with low Km - for use in assay for
 XX creatine.
 XX Disclosure; Page 13-14; 21pp; English.
 XX A novel creatine amidinohydroxylase enzyme has been developed which
 XX catalyses the reaction of creatine with water to form sarcosine and urea,
 XX 18 stable at temperatures of up to 50 degrees Celsius (pH 7.5, 30

CC minutes) and has an optimum temperature of 40-50 degrees Celsius, optimum
 CC pH of 8-9, a Km value for creatine of 3.5-10 mM in a binding assay using
 CC sarcosine oxidase and peroxidase, a molecular weight of 43 kD by SDS-PAGE
 CC and an isoelectric point of 3.5. The present sequence represents creatine
 CC amidinohydroxylase derived from Alcaligenes faecalis strain TE3581 (FERM P-
 CC 14237), which is the wild type creatine amidinohydroxylase to be mutated in
 CC the present invention. The enzyme can be used to determine creatine in a
 CC sample by measuring the absorbance of a dye formed by reacting the sample
 CC with a reagent, comprising the enzyme, sarcosine oxidase and a
 CC composition for detecting hydrogen peroxide, e.g. for diagnosis of
 CC uraemia, chronic nephritis, gigantism and tonic muscular dystrophy. The
 CC enzyme has a lower Km value than prior art creatine amidinohydroxylase (cf.
 CC US 5451520). (Updated on 17-OCT-2003 to standardise OS field)

XX SQ Sequence 404 AA;

Query Match 100.0%; Score 2190; DB 2; Length 404;
 Best Local Similarity 100.0%; Pred. No. 1.5e-214;
 Matches 404; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTDDMLHVKMHNGEKDYSPFSDAEMTRONDVGMMAKNVDAALFTSYHCINYSGL 60
 DB 1 MTDDMLHVKMHNGEKDYSPFSDAEMTRONDVGMMAKNVDAALFTSYHCINYSGL 60

QY 61 YCYFGRKYGWVIDHNNATTISAGIDGQPMRRSFGDNITYTDWRDNFRAVQLTGAK 120
 DB 61 YCYFGRKYGWVIDHNNATTISAGIDGQPMRRSFGDNITYTDWRDNFRAVQLTGAK 120

QY 121 RIGIEFDHNLDPFRQLEALPGVEFVDISQPSMMRTIKSLEOKLIREGARVCVGA 180
 DB 121 RIGIEFDHNLDPFRQLEALPGVEFVDISQPSMMRTIKSLEOKLIREGARVCVGA 180

QY 181 ACAAIKAGVBEHEVAIATTNAMIIRIAKSPPEVLMOTWTWFOGINTDGANPVNRI 240
 DB 181 ACAAIKAGVBEHEVAIATTNAMIIRIAKSPPEVLMOTWTWFOGINTDGANPVNRI 240

QY 241 VQSGDILSLNTPMIFGYTALERTLPCDHVDASLDIWEKNAVHRGELIKPGARCK 300
 DB 241 VQSGDILSLNTPMIFGYTALERTLPCDHVDASLDIWEKNAVHRGELIKPGARCK 300

QY 301 DIAIELNEMRYEMDLKYRSFGYHSGFVLCYHGREAGVELREDIDTELKPGVVSMEP 360
 DB 301 DIAIELNEMRYEMDLKYRSFGYHSGFVLCYHGREAGVELREDIDTELKPGVVSMEP 360

QY 361 MWMLPEGMPGAGGYREHDILIVGEDGAENITGPPFGEHNIIIN 404
 DB 361 MWMLPEGMPGAGGYREHDILIVGEDGAENITGPPFGEHNIIIN 404

RESULT 3
 AAM61905 ID AAM61905 standard; protein; 404 AA.

AC AAM61905;
 DT 16-SEP-1998. (first entry)

DE Stable creatine amidinohydroxylase enzyme.
 KM Createine amidinohydroxylase; mutant; stable; enzyme; diagnostic agent.

OS Alcaligenes faecalis.
 FH Key Location/Qualifiers
 FT Misc-difference 145 /note= "encoded by GAC"

XX JP10174585-A.
 XX 30-JUN-1998.
 XX 17-DEC-1996; 96JP-00337027.

PR	17-DEC-1996;	96GP-00337027.
XX	(TOYM) TOYOB0 KK.	
PA	WPI; 1998-421167/36.	
DR	N-PsDB; AAV35699.	
XX	New creatine amidino-hydrazide used as diagnostic agent - is more stable in neutral buffer than wild type creatine amidino-hydrazide.	
PT	Claim 5; Page 11-13; 14pp; Japanese.	
PS	This represents a stable creatine amidino-hydrazide which is a mutant creatine amidino-hydrazide and has improved long-term stability in a neutral buffer compared to wild type creatine amidino-hydrazide. A recombinant plasmid containing the stable creatine amidino-hydrazide gene can be used to transform a cell for the recombinant production of the enzyme. This stable creatine amidino-hydrazide is useful as a diagnostic agent can be produced commercially	
CC	Sequence 404 AA;	
XX		
SQ		
Query Match	100.0%; Score 2190; DB 2; Length 404;	
Best Local Similarity	100.0%; Pred. No. 1.5e-214;	
Matches	404; Conservative 0; Mismatches 0; Indels 0; Gaps 0	
Qy	1 MTDDMLHWKMHNGEKDYSPPSDEATWRQRNDVRGMAKNVDAALFTSYHCINYYSGWL	60
Dd	1 MTDDMLHWKMHNGEKDYSPPSDEATWRQRNDVRGMAKNVDAALFTSYHCINYYSGWL	60
Qy	61 YCYPERKKGMYVDHNNATTISAGIDGGPWRKSFGDNITYTWRQDNFRAVRQLTTGAK	120
Dd	61 YCYPERKKGMYVDHNNATTISAGIDGGPWRKSFGDNITYTWRQDNFRAVRQLTTGAK	120
Qy	121 RIGIFPDHYNDFRQLEBALPGVEFPDISOPSMRRITKSLEOKLIREGARVCDDVGA	180
Dd	121 RIGIFPDHYNDFRQLEBALPGVEFPDISOPSMRRITKSLEOKLIREGARVCDDVGA	180
Qy	181 ACAAAIKAGVPEHEVAIAITNAMIREIAKSPFEVLMTWTWFOSGINTDGANPVTRRI	240
Dd	181 ACAAAIKAGVPEHEVAIAITNAMIREIAKSPFEVLMTWTWFOSGINTDGANPVTRRI	240
Qy	241 VOSGDIIISLTTPPMIFGYTYTLERLLFCDRVDDASLDIWEKVAVHRGLIELIKGACCK	300
Dd	241 VOSGDIIISLTTPPMIFGYTYTLERLLFCDRVDDASLDIWEKVAVHRGLIELIKGACCK	300
Qy	301 DIAIEINMEYREMDLLKYRSFGYGHSFGVLCHYYGREAGVELREDIDTELKRGVVSNMP	360
Dd	301 DIAIEINMEYREMDLLKYRSFGYGHSFGVLCHYYGREAGVELREDIDTELKRGVVSNMP	360
Qy	361 MVMLPEGMPGAGYREHDILIVGEDGAEINTGFFPPEPHNIIRN	404
Dd	361 MVMLPEGMPGAGYREHDILIVGEDGAEINTGFFPPEPHNIIRN	404
RESULT 4		
AAB09976		
ID	AAB09976 standard; protein; 404 AA.	
XX		
AC	AAB09976;	
DT	19-OCT-2000 (first entry)	
XX		
DE	Alcaligenes sp. creatine amidinohydrolase protein.	
KM	Creatine amidinohydrolase; thermostable; diagnosis; kidney disease; bilirubin.	
OS	Alcaligenes sp.	
PN	WO200031245-A1.	
DD	02-JUN-2000.	

XX	PF	25-NOV-1999;	99W0-JP006583.
XX	PR	25-NOV-1998;	98UP-00334252.
XX	PA	(KIKK) KIKKOMAN CORP.	
XX	PI	Furukawa K, Ichikawa T,	
XX	DR	WPI, 2000-411951/35.	
XX	PT	Highly thermostable creatine amidinohydrolase with optimum pH in weakly	
XX	PT	acidic region, useful in assaying serum or urine creatine for diagnosis	
XX	PT	of e.g. kidney diseases, scarcely affected by bilirubin.	
XX	PS	Claim 3; Page 19-20; 23pp; Japanese.	
XX	CC	This invention describes a novel highly thermostable creatine	
XX	CC	amidinohydrolase (I) with optimum pH in a weakly acidic region, useful in	
XX	CC	assaying serum or urine creatine for diagnosis of e.g. kidney diseases,	
XX	CC	and scarcely affected by bilirubin. The enzyme is produced by a	
XX	CC	transformant <i>Escherichia coli</i> in high efficiency. Due to the enzyme	
XX	CC	having an optimum pH value in the weakly acidic region, it is scarcely	
XX	CC	affected by bilirubin when assaying creatine	
XX	SQ	Sequence 404 AA;	
XX	Query Match	100.0%; Score 2190; DB 3; Length 404;	
XX	Best Local Similarity	100.0%; Pred. No. 1.5e-214;	
XX	Matches	404; Conservative 0; Mismatches 0; Indels 0; Gaps 0	
QY	1	MTDDMLHWKMKNGEKDYSPPSDAEMTRRONDVRGMAKNVDAALFTSYHCINYYSGWL	60
DB	1	MTDDMLHWKMKNGEKDYSPPSDAEMTRRONDVRGMAKNVDAALFTSYHCINYYSGWL	60
QY	61	YCYGRKTKGMYIDHNNAATTISAGIDGGQPMRRSPGDNITTYDWRDNPFRAYROLTTGAK	120
DB	61	YCYGRKTKGMYIDHNNAATTISAGIDGGQPMRRSPGDNITTYDWRDNPFRAYROLTTGAK	120
QY	121	RIGIEFDHVNLDFFRQLEBALPGVEFVDISOPSMWRTTKSLSEOKLREGARVCDVGA	180
DB	121	RIGIEFDHVNLDFFRQLEBALPGVEFVDISOPSMWRTTKSLSEOKLREGARVCDVGA	180
QY	181	ACAAAIKAGVEBEHVAIATTNAMIREIAKSPPEVEMDWTWFOGINTDGAHPVTNRI	240
DB	181	ACAAAIKAGVEBEHVAIATTNAMIREIAKSPPEVEMDWTWFOGINTDGAHPVTNRI	240
QY	241	VQSGDILSLNTPPMIFGYTTLAERTLPCDHVDVDSLDITWENKVAVHRGLELIKRGARCK	300
DB	241	VQSGDILSLNTPPMIFGYTTLAERTLPCDHVDVDSLDITWENKVAVHRGLELIKRGARCK	300
QY	301	DIALELMBMYEPMOLLKYRSFGYHSPFGLCHYYGREAGVELREDIDTELKPGMYVSMPE	360
DB	301	DIALELMBMYEPMOLLKYRSFGYHSPFGLCHYYGREAGVELREDIDTELKPGMYVSMPE	360
QY	361	MVMLPEGMFGAGYREHDLIVGEDGAEITGPPGPEHNIIRN	404
DB	361	MVMLPEGMFGAGYREHDLIVGEDGAEITGPPGPEHNIIRN	404
XX	RESULT 5		
XX	AAU08727	standard; protein; 404 AA.	
XX	AAU08727;		
XX	19-DEC-2001	(first entry)	
XX	DE	Creatine amidinohydrolase polypeptide.	
XX	KW	Creatine amidinohydrolase; water; sarcosine; urea; creatinine; uraemia;	
XX	KW	chronic nephritis; acute nephritis; tonic muscular dystrophy; giantism;	
XX	KW	pigment absorbance.	

XX OS Alcaligenes faecalis.
 XX PN BP132467-A2.
 XX PD 12-SEP-2001.
 XX PF 13-FEB-1997; 2001EP-00113052.
 XX PR 13-FEB-1996; 96GP-00025435.
 XX PR 13-FEB-1997; 97EP-00102270.
 XX PA (TOYM) TOYO BOSEKI KK.
 XX PI Sogabe A, Hattori T, Nishiya Y, Kawamura Y;
 XX DR WPI; 2001-612481/71.
 XX DR N-PSDB; AAS14742.
 XX PT New creatine amidinohydrolase, useful as a routine reagent for clinical
 XX PT tests for determining creatine and creatinine in biological samples,
 XX PT particularly useful in diagnosing diseases such as uremia or chronic
 XX PT nephritis.
 XX PS Disclosure; Page 14-15; 21pp; English.
 XX CC The invention relates to Alcaligenes faecalis creatine amidinohydrolase,
 XX CC which catalyses creatine and water to sarcosine and urea. Creatine
 XX CC amidinohydrolase can be produced by culturing a microorganism producing
 XX CC the protein in a nutrient medium and recovering the protein from the
 XX CC resulting culture. Creatine amidinohydrolase is useful as a routine
 XX CC reagent for clinical tests for determining creatine and creatinine in
 XX CC biological samples. This is particularly useful in diagnosing diseases
 XX CC such as uremia, chronic nephritis, acute nephritis, glaucoma and tonic
 XX CC muscular dystrophy. The presence of creatine in a sample can be
 XX CC determined by measuring an absorbance of a pigment produced by the
 XX CC reaction of a reagent containing creatine amidinohydrolase with the
 XX CC sample. This sequence represents Alcaligenes faecalis creatine
 XX CC amidinohydrolase polypeptide
 XX CC
 XX SQ Sequence 404 AA;
 Query Match 100.0%; Score 2190; DB 4; Length 404;
 Best Local Similarity 100.0%; Pred. No. 1.5e-214;
 Matches 404; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MTDDMLHWKMHNGEKDYSPPSDAEMTRRQNDVRGMANKNVDAALFTSYHCINYYSGWL 60
 DB 1 MTDDMLHWKMHNGEKDYSPPSDAEMTRRQNDVRGMANKNVDAALFTSYHCINYYSGWL 60
 QY 61 YCYFGRKYGWVIDHNNATTISAGIDGQPMWRSPFGDNITYTDMRDNFRAVRQLTTGAK 120
 DB 61 YCYFGRKYGWVIDHNNATTISAGIDGQPMWRSPFGDNITYTDMRDNFRAVRQLTTGAK 120
 QY 121 RIGIEFDHVNLDPRROLEALPGVEFVDISQPSMMRRITKSLBEQKLREGARVCDVGCA 180
 DB 121 RIGIEFDHVNLDPRROLEALPGVEFVDISQPSMMRRITKSLBEQKLREGARVCDVGCA 180
 QY 121 RIGIEFDHVNLDPRROLEALPGVEFVDISQPSMMRRITKSLBEQKLREGARVCDVGCA 180
 DB 121 RIGIEFDHVNLDPRROLEALPGVEFVDISQPSMMRRITKSLBEQKLREGARVCDVGCA 180
 QY 181 ACAAAIKAGVPEHEVAIATTNAMIKEIKSPFEVLMOTMTWFOGINTDGAHNPVTNRI 240
 DB 181 ACAAAIKAGVPEHEVAIATTNAMIKEIKSPFEVLMOTMTWFOGINTDGAHNPVTNRI 240
 QY 241 VOSGDIISLNTFPMIFGYTTALERTLFCDHVDASLDIWEKNVAVHRRGLEIKPGARCK 300
 DB 241 VOSGDIISLNTFPMIFGYTTALERTLFCDHVDASLDIWEKNVAVHRRGLEIKPGARCK 300
 QY 301 DIAIENMYREMBDLKTRSPGYGHSFGVLCYHYGREGAVELREDIDTELKPGMVSMER 360
 DB 301 DIAIENMYREMBDLKTRSPGYGHSFGVLCYHYGREGAVELREDIDTELKPGMVSMER 360
 QY 361 MVMLEPGMPGAGYREHDLIVGEDGAENITGFPFPEHNIIRN 404
 DB 361 MVMLEPGMPGAGYREHDLIVGEDGAENITGFPFPEHNIIRN 404

RESULT 6
 AAM51471
 ID AAM51471 standard; protein, 404 AA.
 XX AC AAM51471;
 XX AC 07-AUG-2003 (revised)
 XX AC 28-JAN-2002 (first entry)
 XX DE Alcaligenes faecalis creatineamidinohydrolase.
 XX KW Alcaligenes faecalis; TE3581; FERM P14237; creatineamidinohydrolase.
 XX OS Alcaligenes faecalis.
 XX OS
 XX FH Key Location/Qualifiers
 XX FH MISC-difference 145
 XX FT MISC-difference /note= "Encoded by GAC"
 XX FT JP2001252088-A.
 XX PN 18-SEP-2001.
 XX PD 16-MAY-1995; 2001JP-00051054.
 XX PF 16-MAY-1995; 95JP-00117283.
 XX PR (TOYM) TOYOBO KK.
 XX PA WPI; 2002-003140/01.
 XX DR N-PSDB; AAI99856.
 XX XX
 XX PS A gene encoding creatineamidinohydrolase.
 XX PT Claim 1; Page 9-10; 11pp; Japanese.
 XX CC The invention relates to Alcaligenes faecalis TE3581 (FERM P-14237)
 XX CC creatineamidinohydrolase and the encoding gene. The gene can be used for
 XX CC the commercial preparation of creatineamidinohydrolase. (Updated on 07-
 XX CC AUG-2003 to correct OS field.)
 XX CC
 XX SQ Sequence 404 AA;
 Query Match 100.0%; Score 2190; DB 5; Length 404;
 Best Local Similarity 100.0%; Pred. No. 1.5e-214;
 Matches 404; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MTDDMLHWKMHNGEKDYSPPSDAEMTRRQNDVRGMANKNVDAALFTSYHCINYYSGWL 60
 DB 1 MTDDMLHWKMHNGEKDYSPPSDAEMTRRQNDVRGMANKNVDAALFTSYHCINYYSGWL 60
 QY 61 YCYFGRKYGWVIDHNNATTISAGIDGQPMWRSPFGDNITYTDMRDNFRAVRQLTTGAK 120
 DB 61 YCYFGRKYGWVIDHNNATTISAGIDGQPMWRSPFGDNITYTDMRDNFRAVRQLTTGAK 120
 QY 121 RIGIEFDHVNLDPRROLEALPGVEFVDISQPSMMRRITKSLBEQKLREGARVCDVGCA 180
 DB 121 RIGIEFDHVNLDPRROLEALPGVEFVDISQPSMMRRITKSLBEQKLREGARVCDVGCA 180
 QY 121 RIGIEFDHVNLDPRROLEALPGVEFVDISQPSMMRRITKSLBEQKLREGARVCDVGCA 180
 DB 121 RIGIEFDHVNLDPRROLEALPGVEFVDISQPSMMRRITKSLBEQKLREGARVCDVGCA 180
 QY 181 ACAAAIKAGVPEHEVAIATTNAMIKEIKSPFEVLMOTMTWFOGINTDGAHNPVTNRI 240
 DB 181 ACAAAIKAGVPEHEVAIATTNAMIKEIKSPFEVLMOTMTWFOGINTDGAHNPVTNRI 240
 QY 241 VOSGDIISLNTFPMIFGYTTALERTLFCDHVDASLDIWEKNVAVHRRGLEIKPGARCK 300
 DB 241 VOSGDIISLNTFPMIFGYTTALERTLFCDHVDASLDIWEKNVAVHRRGLEIKPGARCK 300
 QY 301 DIAIENMYREMBDLKTRSPGYGHSFGVLCYHYGREGAVELREDIDTELKPGMVSMER 360
 DB 301 DIAIENMYREMBDLKTRSPGYGHSFGVLCYHYGREGAVELREDIDTELKPGMVSMER 360

OY 361 MWLPEGMPGAGYREHDLIVGEDGAENITGPPGPEHNIIRN 404
 DB 361 MWLPEGMPGAGYREHDLIVGEDGAENITGPPGPEHNIIRN 404

RESULT 7

ID ABB05660
 ID ABB05660 standard; protein; 404 AA.

AC ABB05660;

DT 29-APR-2002 (first entry)

DE Creatine amidinohydrazase protein SEQ ID NO:1.

KM Creatine amidinohydrazase; enzyme; mutant; stable; clinical diagnosis.

OS Alcaligenes faecalis.

PH Key Location/Qualifiers

FT Misc-difference 145 /note="encoded by GAC"

PN JP2001346594-A.

PD 18-DEC-2001.

PF 17-DEC-1996; 2001JP-00121708.

PR 17-DEC-1996; 96JP-00337027.

PA (TOYM) TOYOBO KK.

DR WPI: 2002-145187/19.

DR N-PSDB; ABA93696.

PT Mutant creatine amidinohydrazase for use in clinical diagnosis has long-term stability in a neutral buffer solution.

PS Claim 5; Page 10-11; 15pp; Japanese.

XX The present invention describes a stable mutant creatine amidinohydrazase having a long-term stability in a neutral buffer solution. Also described are: (1) a creatine amidinohydrazase gene encoding the above stable creatine amidinohydrazase; (2) a gene encoding a mutant creatine amidinohydrazase having a long-term stability in a neutral buffer solution compared to wild type creatine amidinohydrazase, in which at least one gene among those found in a fully defined 1212 nucleotide sequence (see ABA93696) is replaced by another gene; (3) a recombinant plasmid containing a gene encoding the above creatine amidinohydrazase; (4) a cell transformed by the above plasmid; and (5) a method for the preparation of a stable creatine amidinohydrazase in which the above cell is cultured in a medium and creatine amidinohydrazase is collected. The creatine amidinohydrazase is useful as a clinical diagnosing agent. The present sequence represents a creatine amidinohydrazase protein sequence given in the present invention

XX Sequence 404 AA;

Query Match 100.0%; Score 2190; DB 5; Length 404;
 Best Local Similarity 100.0%; Pred. No. 1.5e-214;
 Matches 404; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MTDDMLHYMKHNGEKDYPSPDAEMTRQNDVRGMAKNVDAALFTSYHCINITYSGWL 60
 DB 1 MTDDMLHYMKHNGEKDYPSPDAEMTRQNDVRGMAKNVDAALFTSYHCINITYSGWL 60
 OY 61 YCYFGRKKGWVDHNNATTISAGIDGQPMWRSPGDNITTYDWRDNYFRAVROLTTGAK 120
 DB 61 YCYFGRKKGWVDHNNATTISAGIDGQPMWRSPGDNITTYDWRDNYFRAVROLTTGAK 120
 OY 121 RIGIFDHVNLDFRRQLSEALPGVEFVDISQSMWRRTIKSLSEQKLIREGARVCDVGGA 180
 DB 121 RIGIFDHVNLDFRRQLSEALPGVEFVDISQSMWRRTIKSLSEQKLIREGARVCDVGGA 180

DB 121 RIGIFDHVNLDFRRQLSEALPGVEFVDISQSMWRRTIKSLSEQKLIREGARVCDVGGA 180
 OY 181 ACAAAIKAGVPEHEVAIATTNMIKEIASPPVELMDWTWPGSGINDGANPYTNRI 240
 DB 181 ACAAAIKAGVPEHEVAIATTNMIKEIASPPVELMDWTWPGSGINDGANPYTNRI 240
 OY 241 VOSGDIISLNTFPMIFGYTTALERTLFCDHVDASIDIVEKNVAVRRGLBLIKPGARCK 300
 DB 241 VOSGDIISLNTFPMIFGYTTALERTLFCDHVDASIDIVEKNVAVRRGLBLIKPGARCK 300
 OY 301 DIAIENEMRYEMDILKTRSPFGYHSFGVLCYHGRBAGVEIRBDITELKGMVVSMBP 360
 DB 301 DIAIENEMRYEMDILKTRSPFGYHSFGVLCYHGRBAGVEIRBDITELKGMVVSMBP 360
 OY 361 MWLPEGMPGAGYREHDLIVGEDGAENITGPPGPEHNIIRN 404
 DB 361 MWLPEGMPGAGYREHDLIVGEDGAENITGPPGPEHNIIRN 404

RESULT 8

ID AAR94463
 ID AAR94463 standard; protein; 404 AA.

AC AAR94463;

DT 16-OCT-2003 (revised)

DT 26-SEP-1996 (first entry)

DE Creatine amidinohydrolase.

KM Creatine amidinohydrolase; CAH; kidney; disease; Alcaligenes.

OS Alcaligenes sp; KS-85 FERM BP-4487.

PN DE19536506-A1.

PD 04-APR-1996.

PF 29-SEP-1995; 95DE-01036506.

PR 29-SEP-1994; 94JP-00235737.

PA (KIRK) KIRKMAN CORP.

PI Furukawa K, Ichikawa T, Suzuki M, Koyama Y;

DR WPI: 1996-180805/19.

DR N-PSDB; AAT13291.

PT DNA encoding creatine amidinohydrolase - useful for quantification of creatine to, e.g. diagnose kidney disease.

PS Claim 1; Page 11-13; 18pp; German.

XX CAH is used for quantification of creatine, e.g. to diagnose kidney disease by measuring creatine content of serum or urine. CAH DNA can be inserted into host cells for the prodn. of CAH. CAH can now be produced efficiently without having to add creatine to the culture medium.
 CC (Updated on 16-OCT-2003 to standardise OS field)

XX Sequence 404 AA;

Query Match 98.1%; Score 2149; DB 2; Length 404;
 Best Local Similarity 98.0%; Pred. No. 2.3e-210;
 Matches 396; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 1 MTDDMLHYMKHNGEKDYPSPDAEMTRQNDVRGMAKNVDAALFTSYHCINITYSGWL 60
 DB 1 MTDDMLHYMKHNGEKDYPSPDAEMTRQNDVRGMAKNVDAALFTSYHCINITYSGWL 60
 OY 61 YCYFGRKKGWVDHNNATTISAGIDGQPMWRSPGDNITTYDWRDNYFRAVROLTTGAK 120
 DB 61 YCYFGRKKGWVDHNNATTISAGIDGQPMWRSPGDNITTYDWRDNYFRAVROLTTGAK 120

QY 121 RIGIEFDHVNLDPRQLEALPGVEFVDISQPSMMRTIKSLBEOKLIREGARVCVGA 180
 DB 121 RIGIEFDHVNLDPRQLEALPGVEFVDISQPSMMRTIKSLBEOKLIREGARVCVGA 180
 QY 181 ACAAAIKAGVPEHEVAIATTNAMIIRIAKSPFVELMDTWTWFOGINTDGAHPVTNRI 240
 DB 181 ACAAAIKAGVPEHEVAIATTNAMIIRIAKSPFVELMDTWTWFOGINTDGAHPVTNRI 240
 QY 241 VOSGDILSINTFPMIFGYTTLALERTLFCDHVDDASLDIWEKNVAVHRRGLELIKPGARCK 300
 DB 241 VOSGDILSINTFPMIFGYTTLALERTLFCDHVDDASLDIWEKNVAVHRRGLELIKPGARCK 300
 QY 301 DIAIELNEMRYEMDILKTRSPGYGHSFGVLCYHGREAGVELREDIDTELKPGMVSMMP 360
 DB 301 DIAIELNEMRYEMDILKTRSPGYGHSFGVLCYHGREAGVELREDIDTELKPGMVSMMP 360
 QY 361 MWMLPEGMPGAGGYREHDILIVGEDGAENITGPPGPEHNIIRN 404
 DB 361 MWMLPEGMPGAGGYREHDILIVGEDGAENITGPPGPEHNIIRN 404

RESULT 9

AAB12775
 ID AAB12775 standard; protein; 404 AA.

AC AAB12775;

DT 22-NOV-2000 (first entry)

DE Alcaligenes thermostable creatine amidinohydrolase protein SEQ ID NO:1.

KW Alcaligenes; thermostable creatine amidinohydrolase; kidney disease.

OS Alcaligenes sp.

PN WO200040708-A1.

PD 13-JUL-2000.

PF 28-DEC-1999; 99WO-JP007424.

PR 01-JAN-1999; 99JP-00033359.

PA (KIKK) KIKKOMAN CORP.

PI Furukawa K, Koyama Y, Suzuki M;

XX WPI; 2000-475827/41.

PT Novel thermostable Alcaligenes-derived creatine amidinohydrolase, useful for the diagnosis of kidney diseases and related diseases.

PS Claim 4; Page 20-22; 24pp; Japanese.

CC The present sequence represents a thermostable creatine amidinohydrolase isolated from Alcaligenes sp.. The thermostable creatine amidinohydrolase (I) that hydrolyses 1 mol of creatine to give 1 mol of urea, has a substrate specificity to creatine, has an optimum pH range of 7-8, has a stable pH range of 4-11, has an optimum operating temperature of 45 plus degrees Celsius, is stable at 53 plus degrees Celsius, and has a molecular weight of 92000 Da as determined by the gel filtration method. The enzyme is applicable in diagnosis of kidney diseases and related diseases

SO Sequence 404 AA;

Query Match 98.1%; Score 2149; DB 3; Length 404;
 Best Local Similarity 98.0%; Pred. No. 2.3e-210;
 Matches 396; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 MTDDMLHWKMHNGEVDYSPFSDAEMTRRQNDVRGMKKNVDAALFTSYHCINYYSGWL 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

DB 1 MTDDMLHWKMHNGEVDYSPFSDAEMTRRQNDVRGMKKNVDAALFTSYHCINYYSGWL 60
 QY 61 YCYFGKRYGMVLDHNNATTSAGIDGGQPMRSPGDNITTYTPWRDNPYRAVROLTTGAK 120
 DB 61 YCYFGKRYGMVLDHNNATTSAGIDGGQPMRSPGDNITTYTPWRDNPYRAVROLTTGAK 120
 QY 121 RIGIEFDHVNLDPRQLEALPGVEFVDISQPSMMRTIKSLBEOKLIREGARVCVGA 180
 DB 121 RIGIEFDHVNLDPRQLEALPGVEFVDISQPSMMRTIKSLBEOKLIREGARVCVGA 180
 QY 181 ACAAAIKAGVPEHEVAIATTNAMIIRIAKSPFVELMDTWTWFOGINTDGAHPVTNRI 240
 DB 181 ACAAAIKAGVPEHEVAIATTNAMIIRIAKSPFVELMDTWTWFOGINTDGAHPVTNRI 240
 QY 241 VOSGDILSINTFPMIFGYTTLALERTLFCDHVDDASLDIWEKNVAVHRRGLELIKPGARCK 300
 DB 241 VOSGDILSINTFPMIFGYTTLALERTLFCDHVDDASLDIWEKNVAVHRRGLELIKPGARCK 300
 QY 301 DIAIELNEMRYEMDILKTRSPGYGHSFGVLCYHGREAGVELREDIDTELKPGMVSMMP 360
 DB 301 DIAIELNEMRYEMDILKTRSPGYGHSFGVLCYHGREAGVELREDIDTELKPGMVSMMP 360
 QY 361 MWMLPEGMPGAGGYREHDILIVGEDGAENITGPPGPEHNIIRN 404
 DB 361 MWMLPEGMPGAGGYREHDILIVGEDGAENITGPPGPEHNIIRN 404

RESULT 10

ABR43478
 ID ABR43478 standard; protein; 404 AA.

AC ABR43478;

DT 21-JUL-2003 (first entry)

DE Mutant Erwinia creatinase CTGc2 protein SEQ ID NO:18.

KW Erwinia; creatinase; creatine amidinohydrolase; enzyme; EC 3.5.3.3; Erwinia-type creatinase; creatine; creatinine; uraemia; gigantism; chronic nephritis; acute nephritis; tonic muscular dystrophy; mutant.

OS Erwinia sp.

PN EP1298213-A1.

PD 02-APR-2003.

PF 17-SEP-2002; 2002EP-00020793.

PR 20-SEP-2001; 2001EP-00121780.

PA (HOFF) ROCHE DIAGNOSTICS GMBH.

PI (HOFF) HOFFMANN LA ROCHE & CO AG F.

XX Shao Z, Schumuck R, Kratzsch P, Kenkies J, Weisser H;

XX WPI; 2003-383834/37.

DR N-PSDB; ACC69519.

PT New variant of an Erwinia-type creatinase modified relative to a wild-type creatinase having creatinase activity, useful for determining creatinine and/or creatine concentration in a sample.

PS Example 4; Page 35-36; 51pp; English.

CC The present invention describes a variant of an Erwinia-type creatinase (I) modified relative to a wild-type creatinase having creatinase activity. The variant comprises at least one amino acid substitution at a position of the fully defined 404 amino acid sequence given in ABR43467. CC these are selected from N130, M203, I278, I304 and F395. Creatinase has CC the EC number EC 3.5.3.3 and is also known as creatine amidinohydrolase. CC Also described is a reagent (II) for determining creatine comprising the

CC Erwinia-type creatinase variant. The variant is useful for determining
CC creatinine and/or creatine concentration in a sample. Measuring
CC creatinine and creatine are useful for diagnosing uraemia, chronic
CC nephritis, acute nephritis, gigantism, tonic muscular dystrophy and some
CC other related diseases. The mutant enzymes have improved stability, lower
CC conductivity and/or lower Km-values for creatine: they are much better
CC suited to detection methods for creatine. The present sequence represents
CC a mutant Erwinia creatinase from the present invention

XX Sequence 404 AA;

Query Match

Best Local Similarity 94.3%; Score 2078; DB 6; Length 404;
Matches 380; Conservative 14; Mismatches 10; Indels 0; Gaps 0;

QY 1 MTDDMLHWKMHNGEKDYSFSDAEMTRNDVGMANKNDVALFTSYHCINYSGL 60
DB 1 MTDDMLHWKMHNGEKDYSFSDAEMTRNDVGMANKNDVALFTSYHCINYSGL 60
QY 61 YCFGRKGMVIDHNNATTISAGIDGQPMRSPFGDNITTTDRNDNFYRAVROLTTGAK 120
DB 61 YCFGRKGMVIDHNNATTISAGIDGQPMRSPFGDNITTTDRNDNFYRAVROLTTGAK 120
QY 121 RIGIEFDHVNLDFFRQLEALPGVEFVDISQPSMMNRITKSLEOKLIREGARVCVGA 180
DB 121 RIGIEFDHVNLDFFRQLEALPGVEFVDISQPSMMNRITKSLEOKLIREGARVCVGA 180
QY 181 ACBAATKGVPEHEVALATTNMTREIASSPPELMDTWTWFQSGINDGANNPTNRI 240
DB 181 ACBAATKGVPEHEVALATTNMTREIASSPPELMDTWTWFQSGINDGANNPTNRI 240
QY 241 VOSGDILSLNTPPMIFGYTTALERTLFCDHVDASLDIWEKNVAHRRGLELIKPGARCK 300
DB 241 VOSGDILSLNTPPMIFGYTTALERTLFCDHVDASLDIWEKNVAHRRGLELIKPGARCK 300
QY 301 DIAELNEMRYREMDLLKXRSFGYGSFGVLCHYYGREAGVELREDIDTELKPGMVVSMKP 360
DB 301 DIAELNEMRYREMDLLKXRSFGYGSFGVLCHYYGREAGVELREDIDTELKPGMVVSMKP 360
QY 361 MVMLEPGAGAGYREHDLIVGEDGAENITGPPGPEHNITRN 404
DB 361 MVMLEPGAGAGYREHDLIVGEDGAENITGPPGPEHNITRN 404

RESULT 11

ABR43467 standard; protein; 404 AA.

AC ABR43467;

DT 21-JUL-2003 (first entry)

DE Erwinia sp. (DSM 97-934) creatinase protein SEQ ID NO.2.

KW Erwinia: creatinase; creatine amidinohydrolase; enzyme; EC 3.5.3.3;
KM Erwinia-type creatinase; creatine; creatinine; uraemia; gigantism;
KM chronic nephritis; acute nephritis; tonic muscular dystrophy.

OS Erwinia sp.

PN EP1298213-A1.

PD 02-APR-2003.

PP 17-SEP-2002; 2002EP-00020793.

PR 20-SEP-2001; 2001EP-00121780.

PA (HOFF) ROCHE DIAGNOSTICS GMBH.

PA (HOFF) HOFFMANN LA ROCHE & CO AG F.

PI Shao Z, Schmuck R, Kratzsch P, Kenkles J, Welser H;

DR WPI; 2003-383834/37.
DR N-PSDB; ACC69514.

PT New variant of an Erwinia-type creatinase modified relative to a wild-
PT type creatinase having creatinase activity, useful for determining
PT creatinine and/or creatine concentration in a sample.

PS Claim 1; Page 17-18; 51pp; English.

XX The present invention describes a variant of an Erwinia-type creatinase
CC (1) modified relative to a wild-type creatinase having creatinase
CC activity. The variant comprises at least one amino acid substitution at a
CC position of the fully defined 404 amino acid sequence given in ABR43467;
CC these are selected from N130, M203, I278, I1304 and P395. Creatinase has
CC the EC number EC 3.5.3.3 and is also known as creatine amidinohydrolase.
CC Also described is a reagent (iii) for determining creatine comprising the
CC Erwinia-type creatinase variant. The variant is useful for determining
CC creatinine and/or creatine concentration in a sample. Measuring
CC nephritis, acute nephritis, gigantism, tonic muscular dystrophy and some
CC other related diseases. The mutant enzymes have improved stability, lower
CC conductivity and/or lower Km-values for creatine: they are much better
CC suited to detection methods for creatine. The present sequence represents
CC Erwinia sp. (DSM 97-934) wild-type creatinase from the present invention

XX Sequence 404 AA;

Query Match

Best Local Similarity 94.3%; Score 2078; DB 6; Length 404;
Matches 381; Conservative 14; Mismatches 9; Indels 0; Gaps 0;

QY 1 MTDDMLHWKMHNGEKDYSFSDAEMTRNDVGMANKNDVALFTSYHCINYSGL 60
DB 1 MTDDMLHWKMHNGEKDYSFSDAEMTRNDVGMANKNDVALFTSYHCINYSGL 60
QY 61 YCFGRKGMVIDHNNATTISAGIDGQPMRSPFGDNITTTDRNDNFYRAVROLTTGAK 120
DB 61 YCFGRKGMVIDHNNATTISAGIDGQPMRSPFGDNITTTDRNDNFYRAVROLTTGAK 120
QY 121 RIGIEFDHVNLDFFRQLEALPGVEFVDISQPSMMNRITKSLEOKLIREGARVCVGA 180
DB 121 RIGIEFDHVNLDFFRQLEALPGVEFVDISQPSMMNRITKSLEOKLIREGARVCVGA 180
QY 181 ACBAATKGVPEHEVALATTNMTREIASSPPELMDTWTWFQSGINDGANNPTNRI 240
DB 181 ACBAATKGVPEHEVALATTNMTREIASSPPELMDTWTWFQSGINDGANNPTNRI 240
QY 241 VOSGDILSLNTPPMIFGYTTALERTLFCDHVDASLDIWEKNVAHRRGLELIKPGARCK 300
DB 241 VOSGDILSLNTPPMIFGYTTALERTLFCDHVDASLDIWEKNVAHRRGLELIKPGARCK 300
QY 301 DIAELNEMRYREMDLLKXRSFGYGSFGVLCHYYGREAGVELREDIDTELKPGMVVSMKP 360
DB 301 DIAELNEMRYREMDLLKXRSFGYGSFGVLCHYYGREAGVELREDIDTELKPGMVVSMKP 360
QY 361 MVMLEPGAGAGYREHDLIVGEDGAENITGPPGPEHNITRN 404
DB 361 MVMLEPGAGAGYREHDLIVGEDGAENITGPPGPEHNITRN 404

RESULT 12

ABR43480 standard; protein; 404 AA.

AC ABR43480;

DT 21-JUL-2003 (first entry)

DE Mutant Erwinia creatinase Ctsd7 protein SEQ ID NO.22.

KW Erwinia: creatinase; creatine amidinohydrolase; enzyme; EC 3.5.3.3;
KM Erwinia-type creatinase; creatine; creatinine; uraemia; gigantism;
KM chronic nephritis; acute nephritis; tonic muscular dystrophy; mutant.

```

XX XX Erwinia sp.
XX OS Synthetic.
XX FN EPI298213-A1.
XX PD 02-APR-2003.
XX PF 17-SEP-2002; 2002EP-00020793.
XX PR 20-SEP-2001; 2001EP-00121780.
XX PA (HOPE ) ROCHE DIAGNOSTICS GMBH.
XX PA (HOPE ) HOFMANN LA ROCHE & CO AG F.
XX F1 Shao Z, Schmuck R, Kratzsch P, Kenkies J, Weisser H;
XX XX
XX DR WPI; 2003-383834/37.
XX DR N-PSDB; ACC69521.
XX PT New variant of an Erwinia-type creatinase modified relative to a wild-
XX PT type creatinase having creatinase activity, useful for determining
XX PT creatinine and/or creatine concentration in a sample.
XX PS Example 4; Page 42-44; 51pp; English.
XX CC The present invention describes a variant of an Erwinia-type creatinase
XX CC (I) modified relative to a wild-type creatinase having creatinase
XX CC activity. The variant comprises at least one amino acid substitution at a
XX CC position of the fully defined 404 amino acid sequence given in ABR43467:
XX CC these are selected from N130, M203, I278, I1304 and F395. Creatinase has
XX CC the EC number EC 3.5.3.3 and is also known as creatine amidinohydrolase.
XX CC Also described is a reagent (II) for determining creatine comprising the
XX CC Erwinia-type creatinase variant. The variant is useful for determining
XX CC creatinine and/or creatine concentration in a sample. Measuring
XX CC creatinine and creatine are useful for diagnosing uraemia, chronic
XX CC nephritis, acute nephritis, gigantism, tonic muscular dystrophy and some
XX CC other related diseases. The mutant enzymes have improved stability, lower
XX CC conductivity and/or lower Km-values for creatine: they are much better
XX CC suited to detection methods for creatine. The present sequence represents
XX CC a mutant Erwinia creatinase from the present invention
XX SQ
XX Sequence 404 AA;
XX
Query Match 94.4%; Score 2068; DB 6; Length 404;
Best Local Similarity 93.6%; Pred. No. 4,4e-202;
Matches 378; Conservative 15; Mismatches 11; Indels 0; Gaps 0
QY 1 MTDDDLHWKWKNGEKDYSFSDAEMTRQNDVRGMAMKNDALFTSYHCINYYSGWL 60
DB 1 MTDDDLHWKWKNGEKEYSPFSDAEMTRQSDVRMMAENDVDALFTSYHCINYYSGWL 60
QY 1 XCYFGRKYGWVDHNNATTISAGIOGQGWRRSFGDNITYTDMRDNFYRAVROLTTGAK 120
DB 1 YCYFGRKYGWVDQDHATTISAGIDGQWRRSFGDNITYTDMRDNFQARQLTTPGAR 120
QY 121 RIGIEFDHVNLDFFRLEALPGVEFVDISQSMWRRTIKSLDEOKLREGARVCDVGA 180
DB 121 RIGIEFDHVNLDFFRLEALPGVEFVDISQSMWRRTIKSLDEOKLREGARICDVGA 180
QY 121 RIGIEFDHVNLDFFRLEALPGVEFVDISQSMWRRTIKSLDEOKLREGARICDVGA 180
DB 121 RIGIEFDHVNLDFFRLEALPGVEFVDISQSMWRRTIKSLDEOKLREGARICDVGA 180
QY 181 ACMAALIKGVPHEHEVALATTNMAIREAKSPFVELMDWTWFOGGINDDGHNPTARI 240
DB 181 ACMAALIKGVPHEHEVALATTNMAIREAKSPFVELMDWTWFOGGINDDGHNPTARI 240
QY 241 VOSGILSLNTEPMI FGYVYALERTLCFCHVDASLIDIEKVAHYRRGLIELIKPGARCK 300
DB 241 VOSGILSLNTEPMI FGYVYALERTLCFCHVDASLIDIEKVAHYRRGLIELIKPGARCK 300
QY 301 DIAIELNEMYYREMDLLKYRSFGYGSFGVLCYYGREGAVELREDIDTELKQGMVSMEP 360
DB 301 DIAIELNEMYYREMDLLKYRSFGYGSFGVLCYYGREGAVELREDIDITVLCQGMVSMEP 360
QY 361 MMTLEPGMPGAGCYREHDLIVGEDGAENITGEPPEGEHNITRN 404

```

[illegible]

QY 121 RIGIEFDHVNLDPRROLEALPGVEFVDISQPSMMNRITKSLSEQKLRGARVCDVGGA 180
 DB 121 RIGIEFDHVNLDPRRTLEALPGVEFVDIGQPSMMNRITKSLSEQKLRGARVCDVGGA 180
 QY 181 ACAAIKAGVPEHEVAIATTNAMIKEIAKSPPELMDWTWTFQSGINTDGAHNPVTNRI 240
 DB 181 ACVAIVKAGVPEHEVAIATTNAMIKEIAKSPPELMDWTWTFQSGINTDGAHNPVTNRI 240
 QY 241 VQSGDIISLNTFPMIFGYTTLERTLPCDHVDASLDIWEKNVAVHRGELIKPGARCK 300
 DB 241 VQSGDIISLNTFPMIFGYTTLERTLPCDHVDASLDIWEKNVAVHRGELIKPGARCK 300
 QY 301 DIAELNEMVREWDLTKYRSFGYGHSGYGLCHYYGREAGVELREDIDTELKPGVVSMEP 360
 DB 301 DIAELNEMVREWDLTKYRSFGYGHSGYGLCHYYGREAGVELREDIDTELKPGVVSMEP 360
 QY 361 MVMLEPGAPGAGYREHDLIVGEDGAENITGPPGPEHNITRN 404
 DB 361 MVMLEPGAPGAGYREHDLIVGEDGAENITGPPGPEHNITRN 404

RESULT 14

ABR43477
 ID ABR43477 standard; protein; 404 AA.

AC ABR43477;

DT 21-JUL-2003 (first entry)

DE Mutant Erwinia creatinase CT2m28 protein SEQ ID NO:16.

XX Erwinia; creatinase; creatine amidinohydrolase; enzyme; EC 3.5.3.3;

KW Erwinia-type creatinase; creatine; creatinine; ureamia; gigantism;

KM chronic nephritis; acute nephritis; tonic muscular dystrophy; mutant.

XX Erwinia sp.

OS Synthetic.

PN EPI298213-A1.

XX 02-APR-2003.

PF 17-SEP-2002; 2002EP-00020793.

PR 20-SEP-2001; 2001EP-00121780.

XX (HOFF) ROCHE DIAGNOSTICS GMBH.

PA (HOFF) HOFFMANN LA ROCHE & CO AG F.

XX Shao Z, Schmuck R, Kratzsch P, Kenkiles J, Weisner H;

DR WPI; 2003-383834/37.

XX N-PSDB; ACC69518.

PS Example 4; Page 32-33; 51pp; English.

CC The present invention describes a variant of an Erwinia-type creatinase
 CC (i) modified relative to a wild-type creatinase having creatinase
 CC activity. The variant comprises at least one amino acid substitution at a
 CC position of the fully defined 404 amino acid sequence given in ABR43477;
 CC these are selected from N130, M203, I278, I1304 and P395. Creatinase has
 CC the EC number EC 3.5.3.3 and is also known as creatine amidinohydrolase.
 CC Also described is a reagent (ii) for determining creatine comprising the
 CC Erwinia-type creatinase variant. The variant is useful for determining
 CC creatinine and/or creatine concentration in a sample. Measuring
 CC creatinine and creatine are useful for diagnosing uremia, chronic
 CC nephritis, acute nephritis, gigantism, tonic muscular dystrophy and some
 CC other related diseases. The mutant enzymes have improved stability, lower
 CC conductivity and/or lower Km-values for creatine: they are much better

CC suited to detection methods for creatine. The present sequence represents
 CC a mutant Erwinia creatinase from the present invention

XX SQ Sequence 404 AA;

Query Match 94.3%; Score 2066; DB 6; Length 404;
 Best Local Similarity 93.6%; Pred. No. 7e-202;
 Matches 378; Conservative 16; Mismatches 10; Indels 0; Gaps 0;

QY 1 MTDDMLHMKHNGEKVSPSPDAMETRRQNDVGRGMAKNVDALFTSYHCINYSGL 60
 DB 1 MTDDMLHMKHNGEKVSPSPDAMETRRQNDVGRGMAKNVDALFTSYHCINYSGL 60
 QY 61 YCFGRKKGWYIDHNNATTISAGIDGQPMRSPFCNDITTYDWRDNYFRAVROLTTGAK 120
 DB 61 YCFGRKKGWYIDHNNATTISAGIDGQPMRSPFCNDITTYDWRDNYFRAVROLTTGAK 120
 QY 121 RIGIEFDHVNLDPRROLEALPGVEFVDISQPSMMNRITKSLSEQKLRGARVCDVGGA 180
 DB 121 RIGIEFDHVNLDPRRTLEALPGVEFVDIGQPSMMNRITKSLSEQKLRGARVCDVGGA 180
 QY 181 ACAAIKAGVPEHEVAIATTNAMIKEIAKSPPELMDWTWTFQSGINTDGAHNPVTNRI 240
 DB 181 ACVAIVKAGVPEHEVAIATTNAMIKEIAKSPPELMDWTWTFQSGINTDGAHNPVTNRI 240
 QY 241 VQSGDIISLNTFPMIFGYTTLERTLPCDHVDASLDIWEKNVAVHRGELIKPGARCK 300
 DB 241 VQSGDIISLNTFPMIFGYTTLERTLPCDHVDASLDIWEKNVAVHRGELIKPGARCK 300
 QY 301 DIAELNEMVREWDLTKYRSFGYGHSGYGLCHYYGREAGVELREDIDTELKPGVVSMEP 360
 DB 301 DIAELNEMVREWDLTKYRSFGYGHSGYGLCHYYGREAGVELREDIDTELKPGVVSMEP 360
 QY 361 MVMLEPGAPGAGYREHDLIVGEDGAENITGPPGPEHNITRN 404
 DB 361 MVMLEPGAPGAGYREHDLIVGEDGAENITGPPGPEHNITRN 404

RESULT 15

ABR43475
 ID ABR43475 standard; protein; 404 AA.

AC ABR43475;

DT 21-JUL-2003 (first entry)

DE Mutant Erwinia creatinase CT2m9 protein SEQ ID NO:12.

XX Erwinia; creatinase; creatine amidinohydrolase; enzyme; EC 3.5.3.3;

KW Erwinia-type creatinase; creatine; creatinine; uremia; gigantism;

KM chronic nephritis; acute nephritis; tonic muscular dystrophy; mutant.

XX Erwinia sp.

OS Synthetic.

PN EPI298213-A1.

XX 02-APR-2003.

PF 17-SEP-2002; 2002EP-00020793.

PR 20-SEP-2001; 2001EP-00121780.

XX (HOFF) ROCHE DIAGNOSTICS GMBH.

PA (HOFF) HOFFMANN LA ROCHE & CO AG F.

XX Shao Z, Schmuck R, Kratzsch P, Kenkiles J, Weisner H;

DR WPI; 2003-383834/37.

XX N-PSDB; ACC69518.

PT New variant of an Erwinia-type creatinase modified relative to a wild-
 type creatinase having creatinase activity, useful for determining

PT creatinine and/or creatine concentration in a sample.

XX
PS Example 4; Page 25-26; 51pp; English.

XX
CC The present invention describes a variant of an Erwinia-type creatinase
CC (I) modified relative to a wild-type creatinase having creatinase
CC activity. The variant comprises at least one amino acid substitution at a
CC position of the fully defined 404 amino acid sequence given in ABR43467:
CC these are selected from N130, M203, I278, I1304 and E335. Creatinase has
CC the EC number EC 3.5.3.3 and is also known as creatine amidinohydrolase.
CC Also described is a reagent (II) for determining creatine comprising the
CC Erwinia-type creatinase variant. The variant is useful for determining
CC creatinine and/or creatine concentration in a sample. Measuring
CC creatinine and creatine are useful for diagnosing uraemia, chronic
CC nephritis, acute nephritis, gigantism, tonic muscular dystrophy and some
CC other related diseases. The mutant enzymes have improved stability, lower
CC conductivity and/or lower Km-values for creatine: they are much better
CC suited to detection methods for creatine. The present sequence represents
CC a mutant Erwinia creatinase from the present invention

XX
SQ Sequence 404 AA;

Query Match 94.2%; Score 2064; DB 6; Length 404;

Best Local Similarity 93.6%; Pred. No. 1.1e-201;

Matches 378; Conservative 16; Mismatches 10; Indels 0; Gaps 0;

QY 1 MTDDMLHYMKMNGEKDYSPFSDAEMTRONDVKGMAKNNDALFTSYHCINYSGL 60
DB 1 MTDDMLHYMKMNGEKDYSPFSDAEMTRONDVKGMAKNNDALFTSYHCINYSGL 60
QY 61 YCYFGRKKGWYIDHNNATTISAGIDGQPMWRSPGDNITTYDWRDNYFRAVROLTTGAK 120
DB 61 YCYFGRKKGWYIDHNNATTISAGIDGQPMWRSPGDNITTYDWRDNYFRAVROLTTGAK 120
QY 121 RIGIEFDHVNLDFFRQLEBALPGVEFVDISQPMWRRTIKSLSEQKLIREGARVCDVGA 180
DB 121 RIGIEFDHVNLDFFRQLEBALPGVEFVDISQPMWRRTIKSLSEQKLIREGARVCDVGA 180
QY 181 ACAAAIKAGVPEHEVAIATNNAIIRIAKSPFPELMDTWTFQSGINTDGAHPVTNRI 240
DB 181 ACVAAVKAGVPEHEVAIATNNAIIRIAKSPFPELMDTWTFQSGINTDGAHPVTNRI 240
QY 241 VOSGDILSLNTPPMFGYTTALERTLFCDHVDASLDIWEKNVAVHRGLLEIKPGARCK 300
DB 241 VOSGDILSLNTPPMFGYTTALERTLFCDHVDASLDIWEKNVAVHRGLLEIKPGARCK 300
QY 301 DIAIELNEMYREWDLKTKRSFGYGHSGFVLCYHGREAGVELREDIDTELKPGWVVSMEP 360
DB 301 DIAIELNEMYREWDLKTKRSFGYGHSGFVLCYHGREAGVELREDIDTELKPGWVVSMEP 360
QY 361 MVMLEPGMFGAGGTYREHDLIVGEDGAEINITGPPGPEHNITRN 404
DB 361 MVMLEPGMFGAGGTYREHDLIVGEDGAEINITGPPGPEHNITRN 404

Search completed: July 8, 2005, 01:34:11

Job time : 84 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 8, 2005, 01:01:21 ; Search time 24 Seconds
(without alignments)
1619.649 Million cell updates/sec

Title: US-10-807-228a-1

Perfect score: 2190
Sequence: 1 MTDDMLHVMKHNKGEKDYSP.....DGAENITGPFPGHEHNIIRN 404

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Query length	DB ID	Description
1	1349	61.6	411	2	139809 creatinase (EC 3.5
2	1329.5	60.7	411	2	T44250 creatinase (EC 3.5
3	1048.5	47.9	378	2	UH0134 creatinase (EC 3.5
4	257	11.7	363	1	G69869 Xaa-Pro dipeptidase
5	255	11.6	353	1	C69860 Xaa-Pro dipeptidase
6	248.5	11.3	351	2	E75088 cobalt-dependent p
7	247	11.3	355	2	H83999 Xaa-Pro dipeptidase
8	246.5	11.3	351	1	G71056 probable X-Pro dip
9	232	10.6	406	2	C83867 Xaa-Pro dipeptidase
10	231.5	10.6	365	2	AD1634 X-Pro dipeptidase
11	230.5	10.5	365	2	AB1272 X-Pro dipeptidase
12	229	10.5	361	1	PF5012 hypotrichal prote
13	229	10.5	361	2	A91037 probable peptidase
14	229	10.5	361	2	C85881 probable peptidase
15	227.5	10.4	364	2	C84047 prolidase (proline
16	226.5	10.3	348	2	T46473 X-Pro dipeptidase
17	223.5	10.2	353	2	AB1244 aminopeptidase p h
18	223.5	10.2	353	2	AF1606 aminopeptidase p h
19	222.5	10.2	376	2	A86974 probable cytoplasm
20	222	10.1	356	2	C97158 aminopeptidase p A
21	219.5	10.0	349	2	D75419 proline dipeptidase
22	217	9.9	356	2	F81657 proline dipeptidase
23	216	9.8	372	1	C70658 aminopeptidase p l
24	215.5	9.8	372	2	C86711 aminopeptidase p l
25	210	9.6	376	2	T35868 probable dipeptida
26	208.5	9.5	353	2	D95894 probable hydrolase
27	203	9.3	353	2	A89933 Xaa-Pro dipeptidase
28	197.5	9.0	348	2	H90572 xaa-pro aminopepti
29	197	9.0	356	2	H71496 probable aminopept

30	193.5	8.8	349	2	A72750 probable Xaa-Pro d
31	192.5	8.8	356	1	A71089 probable dipeptida
32	191.5	8.7	359	2	G72425 hypothetical prote
33	191	8.7	362	2	A86831 X-Pro dipeptidase
34	190.5	8.7	360	2	D95185 proline dipeptidase
35	190.5	8.7	392	2	A13141 Xaa-Pro dipeptidase
36	190.5	8.7	360	2	D98052 X-Pro dipeptidase
37	189.5	8.7	360	2	D98052 X-Pro dipeptidase
38	189	8.6	358	2	D97893 X-Pro aminopeptida
39	185.5	8.5	358	2	A97243 xaa-Pro aminopepti
40	183	8.4	353	2	G95021 peptidase M24 fami
41	182	8.3	358	2	H89954 hypothetical prote
42	181.5	8.3	434	2	C95362 probable aminopept
43	180.5	8.2	368	1	S52302 X-Pro dipeptidase
44	175.5	8.0	354	1	C70433 probable X-Pro dip
45	175.5	8.0	355	2	C75134 x-pro aminopeptida

ALIGNMENTS

RESULT 1
I39809
Creatinase (EC 3.5.3.3) - Bacillus sp.
C/Species: Bacillus sp.
C/Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #ext_change 09-Jul-2004
C/Accession: I39809; I39976
R/Suzuki, K.; Sagai, H.; Sugiyama, M.; Imamura, S.
J. Ferment. Bioeng. 76, 77-81, 1993
J. Ferment. Bioeng. 77, 231-234, 1994
A/Title: Molecular cloning and high expression of the Bacillus creatinase gene in Esch
A/Reference number: I39809
A/Accession: I39809
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-411 <RES>
A/Cross-references: UNIPROT:P38487; GB:D14463; NID:G500610; PIDN:BA03358.1; PID:G5006
R/Suzuki, K.; Sagai, H.; Imamura, S.; Sugiyama, M.
J. Ferment. Bioeng. 77, 231-234, 1994
A/Title: Cloning, sequencing, overexpression in Escherichia coli of a sarcosine oxidase
A/Reference number: I39976
A/Accession: I39976
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-46 <RES>
A/Cross-references: GB:D16521; NID:G984787; PIDN:BA03968.1; PID:G840668
C/Genetics:
A/Genes: cre
C/Supfamily: X-Pro aminopeptidase
C/Keywords: hydrolase

Query Match 61.6%; Score 1349; DB 2; Length 411;
Best Local Similarity 63.9%; Pred. No. 1.9e-106;
Matches 253; Conservative 49; Mismatches 90; Indels 4; Gaps 1;

QY	12	HNGEKDYSPEFSDAEMTRQNDVGMMAKNVVAALPTSHVHCINYYSGWLYCYFGRXYGNV	71
DB	15	NGGEKVPKPTFSKXEMTRRNTLRLEWYAKAGIDAVMTSHNTNYSDFLYTSFNRSYALV	74
QY	72	IDHNNATTSAGIDGQFWRSGFNDITYTDRRNNFYAAVQLTT----GAKRIGIEFD	127
DB	75	VYQDGHVTVSANIDGMPWRSGFDENIYTTDKRNNFLYAVKVNNEGFSGRGLGVEND	134
QY	128	HYNIDFPROLEBALPGVEFVDISQPSMMRTIKSLBEQKLREGARVCDVGNAACAAIK	187
DB	135	HMTLDLRQOVODALPNTLELVVVSQAVMGHRMFKSDEIDLINGGARIDIGGAAVEAIR	194
QY	168	AGVPEHYVAIAITNNMIEIKSPFVELIMDTWTFQSGINNDGAMNPVTNNIVYSGDTL	247
DB	195	BGVPEYVALHGTETMVEIKATYPAHLRDTWTFQSGINNDGAMNATSKQDRGDTL	254
QY	248	SLATPMTIFGYVTLALERTLPCDHVDASLDIWEKVAAYHRRGLRIKPGARCKDIAIELN	307
DB	255	SLNCPMTIAGYVTLALERTLFLBEVSDRLLEWEINCKYHRGLRIKPGARCKDIAIELN	314

QY 308 EMTRENDLKKRSFGYCHSGYLCYTGREAGVLEARDIDTELKPGVNVSMEMPVMLPEG 367
 DB 315 EIRREHDLNKRFTGSHSGVLSHYGREGAGLEIRREDIETVLEPGVNVSMEMPVIMLPEG 374
 QY 368 MPGAGYREHDLIVGEGAGENITGPPGPEHNITR 403
 DB 375 EPGAGYREHDLIVISNGTENITKFPFGPEHNITK 410

RESULT 2

T44250
 creatinase (EC 3.5.3.3) [validated] - *Archrobacter* sp. (strain TE1826)

C/Species: *Archrobacter* sp.
 A/Variety: strain TE1826
 C/Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 12-Jul-2004
 C/Accession: T44250
 R/Nishiya, Y.; Toda, A.; Imanaka, T.
 Mol. Gen. Genet. 257, 581-586, 1998
 A/Title: Gene cluster for creatinine degradation in *Archrobacter* sp. TE1826.
 A/Reference number: 222735; MUID:9822334; PMID:9553845
 A/Accession: T44250
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-411 <NIS>
 A/Cross-references: EMBL:AB007122; PDB:BAJ25928.1
 A/Experimental source: strain TE1826
 C/Genetics:
 A/Gene: crea
 A/Function:
 A/Description: EC 3.5.3.3 [validated, MUID:9822334]
 C/Keywords: hydrolase

Query Match 60.7%; Score 1329.5; DB 2; Length 411;
 Best Local Similarity 63.0%; Pred. No. 8.4e-105;
 Matches 250; Conservative 55; Mismatches 87; Indels 5; Gaps 3;

QY 12 HNGEKYSPSDAEMTRONDVGMKANNVDAALFTSYHCINNYSGMLCYFERKXGMV 71
 DB 14 HNGEKYSPSDAEMTRONDVGMKANNVDAALFTSYHCINNYSGMLCYFERKXGMV 73
 QY 72 IDHNNATTISAGIDGQPMRRSFGDNTTYTDMRSDNFYRAVOL--TTGAK--RIGIEFD 127
 DB 74 VTQKAVTVAANIDGMPKRSYDENTVYTDKRDNTFYALQKLEAGVKKALGIEED 133
 QY 128 HYNLDFRQDEALPGVEFVDISQPSMMRTIKSLSEOKLIREGARVCDVGAACAIAIK 187
 DB 134 HVSIDLKRFSDPFNFELVHSQDVMKQRMKSAEIRHIKNGARLADIGGYAVVEAIQ 193
 QY 188 AGPHEHVAATTAAMREIATKSPFVEMDTWMTWFOGINTDGAHNPVTNRIVOSGDI 247
 DB 194 EGPEVEVALAGSVAMREIAKLPQSELRTWTWFGAGINTDGAHNSWATSKVQGEIL 253
 QY 248 SLNTPPMIFGYTALERTLFCDDHYDASLDIWEKNVAVHRRGELIKPGARCKDIATELN 307
 DB 254 SLNTPPMIAGYTLERTLLEESVSDAHLKWEINVEHKGLEILIRPGANCOICALEN 313
 QY 308 EMTRENDLKKRSFGYCHSGVLCYTGREAGVLEARDIDTELKPGVNVSMEMPVMLPEG 367
 DB 314 EMPREHDLVKNRTGYSFGVLSHYGREGAGLEIRREDIETVLEPGVNVSMEMPVIMLPEG 373
 QY 368 MPGAGYREHDLIVGEGAGENITGPPGPEHNITR 403
 DB 374 EPGAGYREHDLIVISNGTENITKFPFGPEHNITK 410

RESULT 3

JH0134
 creatinase (EC 3.5.3.3) - *Flavobacterium* sp.
 N/Alternate names: creatine amidinohydrolase
 C/Species: *Flavobacterium* sp.
 C/Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 09-Jul-2004
 C/Accession: JH0134

R/Koyama, Y.; Kita, S.; Yamamoto-Otake, H.; Suzuki, M.; Nakano, E.

Agric. Biol. Chem. 54, 1453-1457, 1990

A/Title: Cloning and expression of the creatinase gene from *Flavobacterium* sp. U-188

A/Reference number: JH0134; MUID:91103958; PMID:1368564

A/Accession: JH0134

A/Molecule type: DNA

A/Residues: 1-378 <KOY>

A/Cross-references: UNIPROT:P19213

A/Experimental source: strain U-188

C/Comment: This enzyme catalyzes the hydrolysis of creatine to sarcosine and urea.

C/Superfamily: X-Pro aminopeptidase

C/Keywords: hydrolase

Query Match 47.9%; Score 1048.5; DB 2; Length 378;
 Best Local Similarity 59.9%; Pred. No. 5e-81;
 Matches 200; Conservative 42; Mismatches 91; Indels 1; Gaps 1;

QY 5 MHHVMKHNKESDPSDAMTRONDVGMKANNVDAALFTSYHCINNYSGMLCYF 64
 DB 3 MPTLIRNGEKYKSTFSAQEVANRAKLRHIAENIDAIVTSYHNINYSDFLYCSF 62
 QY 65 GRKGYVTDHNNATTISAGIDGQPMRRSFG--DNITYTDMRSDNFYRAVOLTTGARRIG 123
 DB 63 GRPYALVTVDDVISISANIDGQPMRRRTVGTNTIYTWQKRNYPVALQALPRARRIG 122
 QY 124 IERDYNLDFRQDEALPGVEFVDISQPSMMRTIKSLSEOKLIREGARVCDVGAACA 183
 DB 123 IERDYNLDFRQDEALPGVEFVDISQPSMMRTIKSLSEOKLIREGARVCDVGAACA 182
 QY 184 AAIKAGVPEHEVAATTAAMREIATKSPFVEMDTWMTWFOGINTDGAHNPVTNRIVOS 243
 DB 183 EALRDQPEVEVALAGSVAMREIAKLPQSELRTWTWFGAGINTDGAHNSWATSKVQGEIL 242
 QY 244 GDLSLNTPMIFGYTALERTLFCDDHYDASLDIWEKNVAVHRRGELIKPGARCKDIA 303
 DB 243 GDLSLNTPMIFGYTALERTLFCDDHYDASLDIWEKNVAVHRRGELIKPGARCKDIA 302
 QY 304 IERNEMVREHDLIVGEGAGENITGPPGPEHNITR 403
 DB 303 KEINERFLKRDLDQYRTFGYHSGFTLSHYGND 336

RESULT 4

G69869
 Xaa-Pro dipeptidase homolog ykvy - *Bacillus subtilis*

C/Species: *Bacillus subtilis*

C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004

C/Accession: G69869

R/Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bert

A.; Ehrlich, S.D.; Emmerson, P.T.; Eutlian, K.D.; Erlington, J.; Fabret, C.; Ferrati, E

Nature 390, 249-256, 1997

A/Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*

A/Reference number: A69580; MUID:98044033; PMID:9384377

A/Accession: G69869

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Cross-references: UNIPROT:Q31689; GB:Z99111; GB:AL009126; NID:92633699; PDB:1CAB1325

C/Genetics:

A/Gene: ykvy

C/Superfamily: X-Pro aminopeptidase

Query Match 11.7%; Score 257; DB 1; Length 363;
Best Local Similarity 24.9%; Pred. No. 6,1e-14;
Matches 94; Conservative 66; Mismatches 188; Indels 30; Gaps 11;

QY 26 MTRONDVGRMAKNNVDALFTSYHCINYSGLYCYFGKRYG-VIDHNNATTISAGI 84
DB 1 MNRIOF-VSSWLKRGHTAAFIHTKENVFYLLGFTEPHERLMGLFIQEEEPFVCCGM 59
QY 85 DGGQWRSPGDN---ITYTD---WRDNFYRAVROLTTGAKRIGIEFDHVNIDFRROL 137
DB 60 EAGQ-ARNAGMNEITIGYADHENDW---ELIEKALKKKNISIHMLAVKSDISIRADQL 115
QY 138 BEALPGVEFDVDSOPSMWRTIKSLEROKLIREGARVCDVGAACAIAKAVEHEVAI 197
DB 116 KHATGGAQFVSAEELTNGFRILIKODNEIRLKAALADYGVETGATLARGISIVETL- 174
QY 198 ATTAMIR-EIAKSPFPELMDTWTWFOSGINTDGAHPVTNRIVQSGDILSLTFPMIF 256
DB 175 ---AOIEYELKKK--GIQMSFSTWVLFGEKSGCPHNPATATLKKDFVLFDLGVLID 228
QY 257 GYTAALERTLFCDHVDSDLIWENKVAVHRGELIRGARCDIAELNEMTRENDEL 316
DB 229 GYCSDIRTPAYKTINPKQBAIYETVLAERKALIASRPGVIGDLDLTARGLIEKGYG 288
QY 317 KYRSPGYGHSRGLVCHYRGAGVELREDIDTELKPGWVMSBPVWMLPEGMPGAGYRE 376
DB 289 DYFPHRGLHGIGISVHE-----PSMSQANDTLQEGNVYTFIEGIVPE---IGGVRI 339
QY 377 HDLIIVGDEGANITGPP 394
DB 340 EDVHTVKDGAVALTOYP 357

RESULT 5

Xaa-Pro dipeptidase homolog yghT - Bacillus subtilis
C/Species: Bacillus subtilis
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C/Accession: C69960
R/Kuser: F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertet, C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho, A.; Ehrlich, S.D.; Emmerston, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E. Nature 390, 249-256, 1997
A/Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallier, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holteppel, S.; Hosono, S.; Hullo, M.F.; Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Muthers, P.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, Y.; M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, Rieger, M.; Rivolta, C.; Roche, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon, A.; Scher, M.; Schleicher, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron, A.; Winters, P.; Wipet, A.; Yamamoto, H.; Yamane, K.; Yano, K.; Yata, K.; Yoshida, K. A/Authors: Yoshikawa, H.F.; Zumbato, E.; Yoshikawa, H.; Danchin, A.
A/Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis. A/Reference number: A9580; MUID:9804403; PMID:9384377
A/Accession: C69960
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-353 <KUN>
A/Cross-references: UNIPROT:P54518; GB:Z99116; GB:AL009126; NID:92634723; PIDN:CAB14377.
A/Experimental source: strain 168
C/Genetics:
A/Genes: yghT
C/Superfamily: X-Pro aminopeptidase

Query Match 11.6%; Score 255; DB 1; Length 353;
Best Local Similarity 24.9%; Pred. No. 8,6e-14;
Matches 93; Conservative 61; Mismatches 177; Indels 42; Gaps 9;

QY 33 VRGMANKNNVDALFTSYHCINYSGLV---LYCFGRGYAVIDHNNATTISAGIDG 87
DB 6 LRNLFGQLGIDGLITSNTNVRVWTFGTGAGLAVISGDKAAPTDFRYTEQAKVQVGP 65

QY 88 QPMRSPGDNITYTDWRDNFYRAVROLTTGAKRIGIEFDHVNIDFRQLBEALPGVEVF 147
DB 66 EIEHGG-GLSIQTTADYVESF-----GIKLGFGQNSMTGTGTSYSAVISDAILV 115

QY 148 DISOPSMWRTIKSLEROKLIREGARVCDVGAACAIAKAVEHEVAIATTAMIREI 207
DB 116 PVASVEKRLIKSEELIKIEBAALADAFRIHLIFMKGIGSEIAVANELEFMRSG 175
QY 208 AKSPFPELMDTWTWFOSGINTDGAHPVTNRIVQSGDILSLTFPMIFGYTALERTLF 267
DB 176 ADSSSPMIV-----ASGLNSSLPHGVASDKLIESGDLVTLDFGAYKYGCSDIRTVA 229
QY 268 CDHVDSDLIWENKVAVHRGELIRGARCDIAELNEMTRENDELTKRSPGYGHSF 327
DB 230 VQGPSDQLKEIYQVFPDQALGVANHIFPGMTGK---EADALTDHIAAK---GYGDYF 281
QY 328 GVLCHYRGAGVELRED---IDTELKPGWVMSBPVWMLPEGMPGAGYREHDLI 381
DB 282 G---HSTGHLGMEVHESPGLSVRSALIEBGMVTVTEPGIYPE---TGVRIEDDIY 334
QY 382 VGEDGANITGPP 394
DB 335 ITENGNTTITPSP 347

RESULT 6

E75088
cobalt-dependent proline dipeptidase (pep-2) PAB1637 - Pyrococcus abyssi (strain Orsa
C/Species: Pyrococcus abyssi
C/Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C/Accession: E75088
R/Anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A/Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome st
A/Reference number: A75001
A/Accession: E75088
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-351 <KAM>
A/Cross-references: UNIPROT:Q9U2P6; GB:AJ248286; GB:AL096836; NID:95458366; PIDN:CAB50
A/Experimental source: strain Orsay
C/Genetics:
A/Genes: PAB1637
C/Superfamily: X-Pro aminopeptidase

Query Match 11.3%; Score 248.5; DB 2; Length 351;
Best Local Similarity 25.3%; Pred. No. 3e-13;
Matches 92; Conservative 70; Mismatches 161; Indels 41; Gaps 13;

QY 36 WAKARNVDALFTSYHCINYSGLVYCYFGKRYGVIDHNNATTISAGIDGQPMRSPG 95
DB 14 FNNENSIDAVLTIRKNPNIIYLLSGASPLAGY---ILVNDGATLYVPELEYENAKERS-- 68
QY 96 DNTTYTDM-RRDNYFAVROLTTGAKRIGIEFDHVNIDFRQLEBALPGVEFDISQPM 154
DB 69 -KIPVEKFRKDDFYVF---GVKVLGIE-GSLSYFVEDLKEKSKSEFKKVDVYK 122
QY 155 WMTIKSLEROKLIREGARVCDVGAACAIAKAVEHEVAIATTAMIREIAKSPFPV 214
DB 123 EKRIVASDEIKIEKACELADKAVMAALAEVTEGKEREIAKAVEYLMGNAEKRAFD 182
QY 215 ELMDTWTWFOSGINTDGAHPVTNRIVQSGDILSLTFPMIFGYTALERTLFCDHVDA 274
DB 183 -----TIIASGRSLPHGVASDKRIEKGDLVVDLGLVYHNSDITRTVVAGSPNEK 236
QY 275 SLDIWENKVAVHRGELIRGARCDIAELNEMTRENDELTKRSPGYGHSF 327
DB 237 QREIYIVLEAOKKAVEARPGITTK---ELDSIR--NITK--EYGGDYF---IHSL 285
QY 335 GREAGVELRE-----DIDTELKPGWVMSBPVWMLPEGMPGAGYREHDLIIVGDEG 387
DB 286 GRCVGLIEHMEVGVSGYD-EYVLKSGMVTIEGIVIPK---FGVRIEDTIVITKTGA 340

QY 388 ENIT 391
Db 341 RRLT 344

RESULT 7

Xaa-Pro dipeptidase BH2800 (imported) - *Bacillus halodurans* (strain C-125)
C/Species: *Bacillus halodurans*
C/Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C/Accession: H83999
R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A/Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and
A/Reference number: A83650; MUID:20512582; PMID:11058132
A/Accession: H83999
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1355 <STO>
A/Cross-references: UNIPROT:Q9K950; GB:AP001516; GB:BA000004; NID:910175192; PIDN:BA065
A/Experimental source: strain C-125
C/Genetics:
A:Gene: BH2800
C/Superfamily: X-Pro aminopeptidase

Query Match 11.3%; Score 247; DB 2; Length 355;
Best Local Similarity 27.6%; Pred. No. 4, 1e-13;
Matches 78; Conservative 55; Mismatches 122; Indels 28; Gaps 7;

QY 118 GARRIGEFHVNLDPRROLEALPGVEFVDISQPMWMTYKSLEROKLIREGARVCV 177
Db 87 GIGRLGFEKSHVTFE-TYELINKLVSELPVAGLVENIRLIKIDETELQIMQENATADA 145
QY 178 GGAACAARIVAGVEHEVAIATTNAMIETAKSPPELMDTWTFQSGINTDGAHPVT 237
Db 146 AFPHITVIRAGVTERVAVNELPEFRMKQGAESSFDIIV-----ASGRSALPHGVAS 199
QY 238 NRIVQSGDILSTFPMIFGYITALEETLECDHVDASLDIMEKNVAVHRRGLELITPCA 297
Db 200 DKYIEKELVTLDPGAYYKGCSDITRTVAVGDIINDELRIKITYDTVLSAOLKMGEMGIRPGI 259
QY 298 RCDIATLELNMREMDLTKYRSYGYSRFGVLCHYGRAGVARE-----DIDTELK 351
Db 260 TGR---EADALTRDHTAK---GYEYFG--HSGHGLGLEVHEGPGLSMKRAVLK 308
QY 352 PGWVSVMEPMVMLPEGMPAGGYREHDLIVGEDGAENITGFP 394
Db 309 PGWVTVTEPGIYI---SGVGTRIEDDVITTESGNSLTKSP 347

RESULT 8

G71056
Probable X-Pro dipeptidase - *Pyrococcus horikoshii*
C/Species: *Pyrococcus horikoshii*
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 12-Jul-2004
C/Accession: G71056
R/Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekit
M.; Ofuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi
DNA Res. 5, 55-76, 1998
A/Title: Complete sequence and gene organization of the genome of a hyper-thermophilic
A/Reference number: A71000; MUID:98344137; PMID:9679194
A/Accession: G71056
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1351 <KMW>
A/Cross-references: UNIPROT:O58885; GB:AP000005; NID:93236132; PIDN:BA30249.1; PID:9325
A/Experimental source: strain OT3
A/Note: this accession replaces an interim accession for a sequence replaced by GenBank
C/Genetics:
A:Gene: PH1149

Query Match 11.3%; Score 246.5; DB 1; Length 351;

Best Local Similarity 23.4%; Pred. No. 4, 5e-13;
Matches 86; Conservative 75; Mismatches 179; Indels 27; Gaps 8;

QY 26 MTRQNDVGMMAKNVDAALFTSYHCINTYSGMLCYGRKKGMITDHNATITAGID 85
Db 4 MNEKVKKIIIEFMDKNSIDVLIKNPNVYIISGASPLAGY---LITGESATLYYPELE 60
QY 86 GQPMRSPGDNITTYDWR-DNFRVAVQLTGGARIGIEFDHVNLDPRROLEALPGV 144
Db 61 YEMAKES---NIPVEKFKMDFFYALE---GISLIGIE-SSLPGFIEELKKNANIK 112
QY 145 EFDVDSQPMWMTYKSLEROKLIREGARVCVGAACAARIVAGVEHEVAIATTNAMI 204
Db 113 EFKVDVDIRDKRIIESEKIKIIEKACELADKAVMAAIEITEGKEREVAKVEYIMK 172
QY 205 REIYAKSPPELMDTWTFQSGINTDGAHPVTNRIVQSGDILSTFPMIFGYITALE 264
Db 173 MNGAEKPAFD-----TIASGRSALPHGVASDKRIERDVLVIDGALYOHYNSDITR 226
QY 265 TLECDHVDASLDIMEKNVAVHRRGLELITPCARCDIATLELNMREMDLTKYRSFGY 324
Db 227 TLVVGSPNEKQKEIYITVLSAOLKAVESAKPGITTAELDSIANITAEYCGEYFPHSLG 286
QY 325 HSPGVLCHYGRAGVAREHDLITELKPGWVSVMEPMVMLPEGMPAGGYREHDLIVGE 384
Db 287 HGVGLVEHWPVRSQVD-----ETVLRGVITIEPGIYIPK-----IGVRIBDTLITRK 337
QY 385 DGAENIT 391
Db 338 NGSKRLL 344

RESULT 9

C83867
Xaa-Pro dipeptidase BH1739 (imported) - *Bacillus halodurans* (strain C-125)
C/Species: *Bacillus halodurans*
C/Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C/Accession: C83867
R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hi
Nucleic Acids Res. 28, 4317-4331, 2000
A/Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* an
A/Reference number: A83650; MUID:20512582; PMID:11058132
A/Accession: C83867
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1406 <STO>
A/Cross-references: UNIPROT:Q9K950; GB:AP001513; GB:BA000004; NID:910174345; PIDN:BA065
A/Experimental source: strain C-125
C/Genetics:
A:Gene: BH1739

Query Match 10.6%; Score 232; DB 2; Length 406;
Best Local Similarity 24.0%; Pred. No. 9, 3e-12;
Matches 96; Conservative 73; Mismatches 179; Indels 52; Gaps 15;

QY 21 FSDAEMTRQNDVGMMAKNVDAALF---TSYHCINTYSGMLCYFYGRKKGMITDHNAT 77
Db 2 FSLMEIRYRLNRETKGMDBGIVLVSNPSNMYISGYSAMSF-YVHQLLVTLDDPOP 60
QY 78 TTISAGIDGQPMRSPGDN---ITYTD-----WRDNYFAVRQLTTGAKRIGIEF 126
Db 61 LWTGREDASSVYKTTWLBQVITPYPDHYQSETRHPMDFVNIKEIQGKRITIGVEM 120
QY 127 D---HVNLDPRROLEALGVEFVDISQPMWMTYKSLEROKLIREGARVCVGAACA 183
Db 121 DAHYFTGLCYQR-LOGLTNGTFKNAITLINVRLIKSDQELQVWRKAATIAENAKGAY 179
QY 184 AAIKAGVPEHEVAIATTNAMIIR---EIAKSPPELMDTWTFQSGINTDGAHPVTNR 240
Db 180 DTVNVGREDVAAASHAQIKGTAFPGDYSIVM-----LPTGENTISCPHLTWTDRT 234
QY 241 VQSGDILSTFPMIFGYITALEETLEFC---DHVDASLDIMEKNVAVHRRGLELITKPG 296

Db 235 YOEGLDYLTVELIAGCYKRYHVPARTVSLGAPBEHVKEKLAQVIE---GIIHET-LQMIKPG 290
QY 297 ARCDIAIENEMRYEMWDLTKYRSFGYSHSGVGLCHYGRAGYELREDIDTELKPGWV 356
Db 291 VAAEVAATNNQSSISKIGFENKSLGY--SIGLSFPDMGSHWYSLKOGDTLLKPMNTF 348
QY 357 SMEPMVLPKMGPMGAGYREHDIIL---VGEDAENIT 391
Db 349 HL-----MFGI-WYEDYGVETESIRITDGVLELT 378

RESULT 10
AD1634
X-Pro dipeptidase homolog lin1613 [imported] - *Listeria innocua* (strain C11p11262)
C/Species: *Listeria innocua*
C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C/Accession: AD1634
R/Glaeser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurge, O.; Entian, K.D.; Fshhi, H.; Science 294, 849-852, 2001
A/Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapk, G.; Madueno, E.; Maitournam, A.; Mok, C.; Schluter, T.; Simoes, N.; Tietz, A.; Vazquez-Boland, J.A.; Voss, H.; Wehler, A./Title: Comparative genomics of *Listeria* species.
A/Reference number: AB1077; MUID:21537279; PMID:11679669
A/Accession: AD1634
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-365 <GLA>
A/Cross-references: UNIPROT:Q92BD7; GB:AL592022; PIDN:CAC96844.1; PID:gl6411007; GSPDB:C
A/Experimental source: strain C11p11262
C/Genetics:
A/Gene: lin1613
C/Superfamily: X-Pro aminopeptidase

Query Match 10.6%; Score 231.5; DB 2; Length 365;
Best Local Similarity 22.1%; Pred. No. 8.9e-12;
Matches 85; Conservative 69; Mismatches 188; Indels 43; Gaps 9;

QY 26 MTRKQNDVRCGMKANNVDAALFTSYHCINYSGLYCYFGKRYGMVIDHNNATTISA--- 82
Db 1 MEKNIDVLQWMLDQGALEVAFITDPENIAVFGSHSPHERVGLAVFSDSEPLFTPAL 60
QY 83 ---GIDGQPMRBSFGDNITTYTDMRDNFYR---AVROLTTGAKRGIFEDVNDLFR 135
Db 61 EVDVVGGMWTHAYGNDT-----ENPKTIADETIKKAVANPSKAIKGMVSVDRYE 114
QY 136 QLEBALPGVEFVNDISOPSMRRTIKSLBEOKLIREGARVCDVGGAACAAIKAGVPEHEV 195
Db 115 QLAGLPSGSSFFIIEHKIEIRLIKTEALIKELKALADY-----AVQGV-DEI 165
QY 196 AIAITNAMIREIAKSPF-----VELMDTWTFQSGINTDGAHPVNTNRIVOSGDIISL 249
Db 166 AEGKTEA---EIVAKIEYEMKKKGVAMSPDTWVLGKNGALPHGTFGETRIKKGDLVLF 222
QY 250 NTFPMIFGYTALERTLFCDHVDDASLDIMEKNAVHRELELIKRGARCKDAIENEM 309
Db 223 DLGAVHKGCSDTTRVAVAFDITDEOKKIYDTLAEQAVAAVADKVGKIGKASEIDLTPRNI 282
QY 310 YREWMDLKRSFEGYSHSGVGLCHYGRAGYELREDIDTELKPGWVVSMEPMVLPKMG 369
Db 283 IRRAGGDFPRLHGLGLASVHEF-----PSITETNSMELQENMVFTEIRGIV---VP 333
QY 370 GAGGYREHDLIVGEDAENITGFP 394
Db 334 GAVGVRIEDDLVYTKGVQVLTTEFP 358

RESULT 11
AB1272
X-Pro dipeptidase homolog lmo1578 [imported] - *Listeria monocytogenes* (strain EGD-e)
C/Species: *Listeria monocytogenes*
C/Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004

C/Accession: AB1272
R/Glaeser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurge, O.; Entian, K.D.; Fshhi, H.; Jones, L.M.; Karet, U.
Science 294, 849-852, 2001
A/Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapk, G.; Madueno, E.; Maitournam, A.; Mok, C.; Schluter, T.; Simoes, N.; Tietz, A.; Vazquez-Boland, J.A.; Voss, H.; Wehler, A./Title: Comparative genomics of *Listeria* species.
A/Reference number: AB1077; MUID:21537279; PMID:11679669
A/Accession: AB1272
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-365 <GLA>
A/Cross-references: UNIPROT:Q8Y6V3; GB:NC_003210; PIDN:CAC99656.1; PID:gl6411007; GSPDB:C
A/Experimental source: strain EGD-e
C/Genetics:
A/Gene: lmo1578
C/Superfamily: X-Pro aminopeptidase

Query Match 10.5%; Score 230.5; DB 2; Length 365;
Best Local Similarity 22.1%; Pred. No. 1.1e-11;
Matches 85; Conservative 67; Mismatches 190; Indels 43; Gaps 9;

QY 26 MTRKQNDVRCGMKANNVDAALFTSYHCINYSGLYCYFGKRYGMVIDHNNATTISA--- 82
Db 1 MEKNIDVLQWMLDQGALEVAFITDPENIAVFGSHSPHERVGLAVFSDSEPLFTPAL 60
QY 83 ---GIDGQPMRBSFGDNITTYTDMRDNFYR---AVROLTTGAKRGIFEDVNDLFR 135
Db 61 EVDVVGGMWTHAYGNDT-----ENPKTIADETIKKAVANPSKAIKGMVSVDRYE 114
QY 136 QLEBALPGVEFVNDISOPSMRRTIKSLBEOKLIREGARVCDVGGAACAAIKAGVPEHEV 195
Db 115 QLAGLPSGSSFFIIEHKIEIRLIKTEALIKELKALADY-----AVQGV-DEI 165
QY 196 AIAITNAMIREIAKSPF-----VELMDTWTFQSGINTDGAHPVNTNRIVOSGDIISL 249
Db 166 AEGKTEA---EIVAKIEYEMKKKGVAMSPDTWVLGKNGALPHGTFGETRIKKGDLVLF 222
QY 250 NTFPMIFGYTALERTLFCDHVDDASLDIMEKNAVHRELELIKRGARCKDAIENEM 309
Db 223 DLGAVHKGCSDTTRVAVAFDITDEOKKIYDTLAEQAVAAVADKVGKIGKASEIDLTPRNI 282
QY 310 YREWMDLKRSFEGYSHSGVGLCHYGRAGYELREDIDTELKPGWVVSMEPMVLPKMG 369
Db 283 IRRAGGDFPRLHGLGLASVHEF-----PSITETNSMELQENMVFTEIRGIV---VP 333
QY 370 GAGGYREHDLIVGEDAENITGFP 394
Db 334 GAVGVRIEDDLVYTKGVQVLTTEFP 358

RESULT 12
F65012
hypothetical protein b2385 - *Escherichia coli* (strain K-12)
C/Species: *Escherichia coli*
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C/Accession: F65012
R/Blatner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A/Title: The complete genome sequence of *Escherichia coli* K-12.
A/Reference number: A64720; MUID:97426617; PMID:9278503
A/Accession: F65012
A/Status: preliminary
A/Molecule type: DNA
A/Status: preliminary
A/Molecule type: DNA
A/Cross-references: UNIPROT:P76524; GB:AE000326; GB:U00096; NID:gl1788718; PIDN:AACT544
A/Experimental source: strain K-12, substrain MG1655
C/Superfamily: X-Pro aminopeptidase

Query Match 10.5%; Score 229; DB 1; Length 361;
Best Local Similarity 23.9%; Pred. No. 1.4e-11;

Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: C84047
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-364 <STO>
A:Cross-references: UNIPROT:Q9K828; GB:AP001518; GB:BA000004; NID:G10175792; PIDN:BA006
A:Experimental source: strain C-125
C:Genetics:
A:Gene: pepQ
C:Superfamily: X-Pro aminopeptidase

Query Match 10.4%; Score 227.5; DB 2; Length 364;
Best Local Similarity 23.6%; Pred. No. 1.9e-11;
Matches 90; Conservative 65; Mismatches 171; Indels 55; Gaps 13;

```
QY 36 WMAKNNVDAALFTSYHCINYSGLYC-----YGRKYGVNIDHNNATTI--SNG 83
   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 10 WLKNDHSMVAFIQDTSIFYLTGF-YCDPHERLVSLLPFAAPCLICPMETSIVKEAG 68
   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 84 IDGGPWRSSFGDNTTITWRDNPYRAVROLTGAKRI---GIEFDHNLDFRROLEE 139
   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 69 -----WT---GEILGYSD--IEDPMLVROAVERKNAVLTSQIVEAPRLTYARVQALQD 117
   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 140 ALPGVEFVDISQPSMMRTIKSLEROKLIREGARVODVGAAACAAIKAGVPEHEVALAT 199
   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 118 AFPFAIKLIDGEPFLMELKQSAKEITTIKENAALADYGEVGVQRIQGRSEIETLALI 177
   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 200 TNAMIREIAKSEPFVELMDTWTWFOGINTDGAHNPVTNRIVOSGDILSLNTPMIFGY 259
   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 178 EVELKRGKGVDMSEFGLTV-----LSGDQSANPHGNPGORTIKKGFVLFDLGVLDGYC 231
   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 260 TALERTLCDFRDASLDIWEKQNAVHRRGLELIKPGARCKDIAIEINEMTREMDLKKR 319
   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 232 SDITRTVAFHVTDOODIYETVRKAQQAALDACRPGVEIRTIQIARTIITE----- 284
   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 320 SFGVGSFGLCHYYGRSAGVELRS---DIDTE-LKPGWVVSMEPMVMLPEGMPGAGG 373
   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 285 -AGYGDYFP--HRIGHGMEVHELPSLNETNTDRLQKQNVFTIEPGIYL---PSIGG 336
   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
QY 374 YREHDILIVGEDGAEINITGFP 394
   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 337 VRIEDDVVITBEDGYQTLTNYP 357
   |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
```

Search completed: July 8, 2005, 01:36:11
Job time : 25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OW protein - protein search, using sw model

Run on: July 7, 2005, 23:34:26 ; Search time 84 Seconds

(without alignments)
2462.857 Million cell updates/sec

Title: US-10-807-228a-1

Perfect score: 2190

Sequence: 1 MTDDMLHVMKHNKGEKDYSP.....DGAENITGFPFGPHNIIIRN 404

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt 03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Fred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2190	100.0	404	2	Q9RHU9
2	1712	78.2	402	2	Q7SINS
3	1386.5	63.3	403	2	Q9R9T5
4	1376.5	62.9	403	2	Q9WWZ1
5	1350.5	61.7	403	2	Q88GQ1
6	1349.5	61.6	403	1	CREA_FIASU
7	1349	61.6	410	1	CREA_BACB0
8	1347.5	61.5	403	1	CREA_PSEBU
9	1329.5	60.7	411	2	O6I182
10	1279.5	12.8	376	2	Q8EMJ3
11	274.5	12.5	364	2	O6SKC1
12	273.5	12.5	357	2	O67N93
13	270	12.3	353	2	O65HH3
14	261.5	11.9	365	2	Q7ZC0
15	261.5	11.9	365	2	Q6HCR7
16	258.5	11.8	365	2	O817E1
17	257	11.7	363	2	O31689
18	255.5	11.7	353	2	Q730Z5
19	255.5	11.7	355	2	Q74BM0
20	255.5	11.7	365	2	O633T3
21	255	11.6	353	1	YQHT_BACSU
22	250.5	11.4	365	2	O81KX5
23	248.5	11.3	351	2	Q9UZP6
24	248.5	11.3	353	2	O8EQ27
25	247.5	11.3	353	2	O81M33
26	247.5	11.3	353	2	O6HMD7
27	247	11.3	355	2	Q9K950
28	246.5	11.3	351	1	PEPO_PYRHO
29	245.5	11.2	353	2	O634Y6
30	245.5	11.2	360	2	Q9HJD2
31	243	11.1	360	2	Q9HJD2

32	241.5	11.0	353	2	O818P9	O818P9 bacillus ce
33	235.5	10.8	354	2	O836X1	O836X1 enterococcu
34	235.5	10.8	365	2	O71Z89	O71Z89 listeria mo
35	234	10.7	421	2	O67R80	O67R80 symbiodace
36	233.5	10.7	356	2	O636F0	O636F0 bacillus ce
37	233	10.6	355	2	O8RAE3	O8RAE3 thermomane
38	232.5	10.6	356	2	O81WG2	O81WG2 bacillus an
39	232	10.6	406	2	O9KC35	O9KC35 bacillus ha
40	231.5	10.6	365	2	O92BD7	O92BD7 listeria in
41	230.5	10.5	365	2	O8Y6V3	O8Y6V3 listeria mo
42	230	10.5	356	2	O819T2	O819T2 bacillus ce
43	229.5	10.5	356	2	O6HER9	O6HER9 bacillus th
44	229	10.5	361	1	YPDF_ECOLI	P76524 escherichia
45	229	10.5	361	2	O8XBP9	O8XBP9 escherichia

ALIGNMENTS

RESULT 1	Q9RHU9	PRELIMINARY	PRT	404 AA.
AC	Q9RHU9			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-JUN-2003 (TREMBLrel. 24, Last annotation update)			
DE	Creatine amidinohydrolase.			
OS	Alcaligenes sp.			
OC	Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;			
OC	Alcaligenaceae; Alcaligenes.			
OX	NCBI_Taxid=512;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=KS-85;			
RA	Furukawa K., Ichikawa T., Koyama Y., Suzuki M.;			
RL	Submitted (Aug-1998) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AB016788; BAA88830.1; -			
DR	HSSP; P38488; 1CHM.			
DR	GO; GO:0016980; F:creatine activity; IEA.			
DR	GO; GO:0016787; F:hydrolase activity; IEA.			
DR	GO; GO:0008235; F:metallopeptidase activity; IEA.			
DR	GO; GO:0006600; P:creatine metabolism; IEA.			
DR	GO; GO:0006508; P:proteolysis and peptidolysis; IEA.			
DR	InterPro; IPR000587; Creatinase.			
DR	InterPro; IPR000994; Peptidase_M24.			
DR	Pfam; PF01321; Creatinase_N; 1.			
DR	Pfam; PF00557; Peptidase_M24; 1.			
KW	Hydrolase.			
SC	SEQUENCE 404 AA; 46032 MW; 82A941891A1ABE79 CRC64;			
Query Match	100.0%; Score 2190; DB 2; Length 404;			
Best Local Similarity	100.0%; Pred. No. 3; 1e-168;			
Matches 404; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	1	MTDDMLHVMKHNKGEKDYSPSPDAEMTRRONDVAGMAKNVDAALFTSYHCINYSGWL	60	O818P9 bacillus ce
DB	1	MTDDMLHVMKHNKGEKDYSPSPDAEMTRRONDVAGMAKNVDAALFTSYHCINYSGWL	60	O818P9 bacillus ce
QY	61	YCYGGRKYGVNIDHNNATTSAGIDGQPMRRSEFNDITTYDWRDNFRAVROLITGAK	120	O818P9 bacillus ce
DB	61	YCYGGRKYGVNIDHNNATTSAGIDGQPMRRSEFNDITTYDWRDNFRAVROLITGAK	120	O818P9 bacillus ce
QY	121	RIGIEPQVNLDRRQLEALPGVEFDISQPSMMWATISLEOKLIREGARCDVGA	180	O818P9 bacillus ce
DB	121	RIGIEPQVNLDRRQLEALPGVEFDISQPSMMWATISLEOKLIREGARCDVGA	180	O818P9 bacillus ce
QY	181	ACAAATKAGVBEHVAATATNAMIREFAKSPFVELADDTWPOSGINTDGAHPVTR	240	O818P9 bacillus ce
DB	181	ACAAATKAGVBEHVAATATNAMIREFAKSPFVELADDTWPOSGINTDGAHPVTR	240	O818P9 bacillus ce
QY	241	VQSGDILSLNTFPPIFGYTTALERTLLFCDHVDASLDIWEKNVAVHRGELIRGARCK	300	O818P9 bacillus ce
DB	241	VQSGDILSLNTFPPIFGYTTALERTLLFCDHVDASLDIWEKNVAVHRGELIRGARCK	300	O818P9 bacillus ce

QY 301 DIATLNMETREMDLTKYRSPGYGHSFGVLCYYYGREAGVELREDIDTELKPGWVSMEP 360
 DB 301 DIATLNMETREMDLTKYRSPGYGHSFGVLCYYYGREAGVELREDIDTELKPGWVSMEP 360
 QY 361 MVMLEPGMPGAGYREHDLIVGEDGAMNITGPFGEHNIIRN 404
 DB 361 MVMLEPGMPGAGYREHDLIVGEDGAMNITGPFGEHNIIRN 404

RESULT 2

07SIBS PRELIMINARY; PRT; 402 AA.
 ID 07SIBS
 AC 07SIBS;
 DT 01-MAR-2004 (Tremblrel. 26, Created)
 DT 01-MAR-2004 (Tremblrel. 26, Last sequence update)
 DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
 DE Hydrolyase.
 OS Actinobacillus.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
 OC Pasteurellaceae.
 NCBI_TaxID=713;
 RN [1]
 RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).
 RX PubMed:12136144; DOI=10.1107/S0907444902010156;
 RA Padmanabhan B., Paehler A., Horikoshi M.;
 RL "Structure of creatine amidohydrolyase from Actinobacillus."
 RL Acta Crystallogr. D Biol. Crystallogr. 58:1322-1328 (2002).
 DR PDB; 1KPO; X-ray; A/B=1-402.
 DR GO; GO:0016980; F:creatinase activity; IEA.
 DR GO; GO:0008235; F:metalloxopeptidase activity; IEA.
 DR GO; GO:0006600; P:creatin metabolism; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR000587; Creatinase.
 DR InterPro; IPR000994; Peptidase_M24.
 DR Pfam; PF01321; Creatinase N; 1.
 DR Pfam; PF00557; Peptidase_M24; 1.
 DR 3D-structure; Hydrolyase.
 SQ SEQUENCE 402 AA; 45872 MW; 2F6B95BA9038B5 CRC64;

Query Match 78.2%; Score 1712; DB 2; Length 402;

Best Local Similarity 75.2%; Pred. No. 1,2e-129;
 Matches 300; Conservative 50; Mismatches 49; Indels 0; Gaps 0;

QY 5 MLHMKHNGKXSPSPDAEMTRONDVGMKKNVDALEFSSYCIYSGMLCYCF 64
 DB 3 MIZTKYHNGZKKTTPSSZAMTRZBLRAMMKSSIDAFLFSYHNINYSGLCYCF 62
 QY 65 GRKXGMYIDHNNATTISAGIDGQPMRRSPGDNITTYDMRDNFYRAVROLTGAKRIGI 124
 DB 63 GRKXAZIYBVKAVTISKIGIDGMPRRSROBNIVTDMKRDNFSYAVKGLVGAOKIGI 122
 QY 125 EFDHVIDFRROLEALPGVEFYDISOPSMWRTIKSLEROKLIREGARVCDVGAACAA 184
 DB 123 EHDHVTLBHRRLZLKALPGTEFYDVVGZPVWVZRVYKSSSEZBLIRZGARISDIGAATAA 182
 QY 185 AIAKGVENHVAATTAAMREIKSPFVELMDTWTFOSGINTGDANHPVNRIVOSG 244
 DB 183 AIAKGVENHVAATTAAMREIKSPFVELMDTWTFOSGINTGDANHPVNRIVOSG 242
 QY 245 DIISLNTFPMIFGYTALERTLFCDHVDASLDIWEKNVAVHRGELIRPGARCKOIAI 304
 DB 243 DIISLNTFPMIFGYTALERTLFLZZVBDASLZIWZKOTAVHRGELIRPGARCKOIAS 302
 QY 305 ELNEMTREMDDLTKYRSPGYGHSFGVLCYYYGREAGVELREDIDTELKPGWVSMEP 364
 DB 303 ELNEMTREMDDLTKYRSPGYGHSFGVLCYYYGREAGVELREDIDTELKPGWVSMEP 362
 QY 365 PEGPAGAGYREHDLIVGEDGAMNITGPFGEHNIIR 403
 DB 363 PEGPAGAGYREHDLIVGEDGAMNITGPFGEHNIIR 401

RESULT 3

09R9T5 PRELIMINARY; PRT; 403 AA.
 ID 09R9T5
 AC 09R9T5;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
 DE Creatinase.
 GN Name-cln;
 OS Pseudomonas putida.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 NCBI_TaxID=303;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RS65;
 RA Liu W.-H., Tang T.-Y., Wen C.-J., Lin Y.-C.;
 RT "Cloning, Sequencing, and Expression of the Creatinase Gene from
 RT Pseudomonas putida RS65 in Escherichia coli."
 RL Food Sci. Agric. Chem. 1:115-121 (1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RS65;
 RA Tang T.-Y., Liu W.-H.;
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF170566; AAD52565.4; -
 DR HSSP; P38488; 1CHM.
 DR GO; GO:0016980; F:creatinase activity; IEA.
 DR GO; GO:0016787; F:hydrolase activity; IEA.
 DR GO; GO:0008235; F:metalloxopeptidase activity; IEA.
 DR GO; GO:0006600; P:creatin metabolism; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR000587; Creatinase.
 DR InterPro; IPR000834; Peptidase_M14.
 DR InterPro; IPR000994; Peptidase_M24.
 DR Pfam; PF01321; Creatinase N; 1.
 DR Pfam; PF00557; Peptidase_M24; 1.
 DR POSITE; PS00133; CARBOXYPEPT_ZN_2; UNKNOWN_1.
 KW Hydrolyase.
 SQ SEQUENCE 403 AA; 45680 MW; D082072C692C1A9A CRC64;

Query Match 63.3%; Score 1386.5; DB 2; Length 403;

Best Local Similarity 66.1%; Pred. No. 2,2e-103;
 Matches 259; Conservative 48; Mismatches 84; Indels 1; Gaps 1;

QY 13 NGKXKXSPSPDAEMTRONDVGMKKNVDALEFSSYCIYSGMLCYCFGRKXGMYI 72
 DB 11 NGKXKXSPSPDAEMTRONDVGMKKNVDALEFSSYCIYSGMLCYCFGRKXGMYI 70
 QY 73 DHNNATTISAGIDGQPMRRSPGDNITTYDMRDNFYRAVROLTGAKRIGIEFDHVN 131
 DB 71 TOEAVVISANIDGQPMRRSPGDNITTYDMRDNFYRAVROLTGAKRIGIEFDHVN 130
 QY 132 DFRROLEALPGVEFYDISOPSMWRTIKSLEROKLIREGARVCDVGAACAAIAKAVP 191
 DB 131 MNRDKLASRYPOLELVDAIPCKRMKMIKSAEHAHILROGARVADIGAAVVALDOVP 190
 QY 192 EHDHVAATTAAMREIKSPFVELMDTWTFOSGINTGDANHPVNRIVOSGDIISLNT 251
 DB 191 EHDHVAATTAAMREIKSPFVELMDTWTFOSGINTGDANHPVNRIVOSGDIISLNT 250
 QY 252 FPMIFGYTALERTLFCDHVDASLDIWEKNVAVHRGELIRPGARCKOIAIENEMYR 311
 DB 251 FPMIFGYTALERTLFLDCHSDHLMWENYVHAGLEIVPGRKCSIALQALNEIFL 310
 QY 312 EMDLTKYRSPGYGHSFGVLCYYYGREAGVELREDIDTELKPGWVSMEP 371
 DB 311 EMDLTKYRSPGYGHSFGVLCYYYGREAGVELREDIDTELKPGWVSMEP 370
 QY 372 GGYREHDLIVGEDGAMNITGPFGEHNIIR 403
 DB 371 GGYREHDLIVGEDGAMNITGPFGEHNIIR 402

```

RESULT 4
Q9WMZ1 PRELIMINARY: PRT: 403 AA.
AC Q9WMZ1,
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Creatinase (EC 3.5.3.3).
OS Pseudomonas putida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NTU-8;
RX MEDLINE=99247136; PubMed=10230521; DOI=10.1023/A:1018705831622;
RA Hong M.C.; Chang J.C.; Wu M.L.; Chang M.C.;
RT "Expression and export of Pseudomonas putida NTU-8 creatinase by
RT Escherichia coli using the chitinase signal sequence of Aeromonas
RT hydrophila."
RL Biochem. Genet. 36:407-415(1998).
DR EMBL; AF072304; AAD37463.1; -.
DR HSSP; P38488; 1CM.
DR GO; GO:0016980; F:creatinase activity; IEA.
DR GO; GO:0016787; F:hydrolyase activity; IEA.
DR GO; GO:0008235; F:metallooxopeptidase activity; IEA.
DR GO; GO:0006600; P:creatinase metabolism; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000587; Creatinase.
DR InterPro; IPR000834; Peptidase_M14.
DR InterPro; IPR000994; Peptidase_M24.
DR Pfam; PF01321; Creatinase_N; 1.
DR Pfam; PF00557; Peptidase_M24; 1.
DR PROSITE; PS00133; CARBOXYPEPT_ZN_2; UNKNOWN_1.
KM Hydrolyase.
SQ SEQUENCE 403 AA; 45691 MW; 0790A5634E7F0E2 CRC64;

Query Match 62.9%; Score 1376.5; DB 2; Length 403;
Best Local Similarity 65.8%; Pred. No. 1.4e-102;
Matches 258; Conservative 48; Mismatches 85; Indels 1; Gaps 1;

QY 13 NGEKDYSPSDAEMTRRQNDVRCMAKKNVDAALFTSYHCINYYSGMLYCYGKRYGVI 72
DB 11 NGEKDYSPSDAEMTRRQNDVRCMAKKNVDAALFTSYHCINYYSGMLYCYGKRYGVI 70
QY 73 DHNNATTISAGIDGGQPMWRSSFG-DNITYTDMRRDNFRAVROLTTGAKRIGIEFDHVL 131
DB 71 TQEVAVSISANIDGGQPMWRSSFG-DNITYTDMRRDNFRAVROLTTGAKRIGIEFDHVL 130
QY 132 DFRQLEALPGVEFVDISQPSMMKRTIKSLBEQKLIBGARVCDVGAACAATKAGVP 191
DB 131 MNRKLSLRYPQAEIVDIAACMRKRMKSAEHHAIIRQGARVADIGAAVVEALRDVP 190
QY 192 EHEVAATTNMIREIAKSPFVELMDTWTFQSGINTDGAHPNTNRIYOSGLISLNT 251
DB 191 EYEVAAHTQAMVREIARTYPSSELDMDTWTFQSGINTDGAHPNTNRIYOSGLISLNT 250
QY 252 FPMIFGYTALERTLFCDHVDASLDIWEKNVAVHRGELIKPGARKDIAIENEMYR 311
DB 251 FPMIFGYTALERTLFCDHVDASLDIWEKNVAVHRGELIKPGARKDIAIENEMYR 310
QY 312 EMDLTKRSFGYGSFGVLCHYRGREAGVELREDIDTBLKPGKVTVMPEPMVNLPGMGFA 371
DB 311 EMDLTKRSFGYGSFGVLCHYRGREAGVELREDIDTBLKPGKVTVMPEPMVNLPGMGFA 370
QY 372 GGYREHDLIIYVGEAGENITGPPGGEHNIR 403
DB 371 GGYREHDLIIYVGEAGENITGPPGGEHNIR 402

```

```

ID Q08GQ1 PRELIMINARY: PRT: 403 AA.
AC Q08GQ1,
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Creatinase.
GN OrderedLocustNames=PP3667;
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160486;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22423060; PubMed=12534463;
RA Nelson K.B.; Weinel C.; Paulsen I.T.; Dodson R.J.; Hilbert H.;
RA Martins dos Santos V.A.P.; Fouts D.E.; Gill S.R.; Pop M.; Holmes M.;
RA Brinkac L.M.; Beanan M.J.; Deboy R.T.; Daugherty S.C.; Kolonay J.F.;
RA Madupu R.; Nelson W.C.; White O.; Peterson J.D.; Khouri H.M.;
RA Hance I.; Chits Lee P.; Holtzapfle E.K.; Scanlan D.; Tran K.;
RA Moazzar A.; Utecherback T.R.; Rizzo M.; Lee K.; Kosack D.; Moestl D.;
RA Wedler H.; Lauber J.; Stjepandic D.; Hobeisel J.; Straetz M.; Helm S.;
RA Kiewitz C.; Eisen J.A.; Timmis K.N.; Duesterhoeft A.; Tsemmler B.;
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
RT metabolically versatile Pseudomonas putida KT2440."
RL Environ. Microbiol. 4:799-808(2002).
DR EMBL; AE016787; AAN69267.1; -.
DR HSSP; P38488; 1CM.
DR TIGR; PP3667; -.
DR GO; GO:0016980; F:creatinase activity; IEA.
DR GO; GO:0016787; F:hydrolyase activity; IEA.
DR GO; GO:0008235; F:metallooxopeptidase activity; IEA.
DR GO; GO:0006600; P:creatinase metabolism; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000587; Creatinase.
DR InterPro; IPR000994; Peptidase_M24.
DR Pfam; PF01321; Creatinase_N; 1.
DR Pfam; PF00557; Peptidase_M24; 1.
KM Complete proteome; Hydrolyase.
SQ SEQUENCE 403 AA; 45633 MW; 7946470E04727B80 CRC64;

Query Match 61.7%; Score 1350.5; DB 2; Length 403;
Best Local Similarity 63.2%; Pred. No. 1.7e-100;
Matches 253; Conservative 50; Mismatches 96; Indels 1; Gaps 1;

QY 5 MLHVMKHNKGEKDYSPSDAEMTRRQNDVRCMAKKNVDAALFTSYHCINYYSGMLYCYF 64
DB 3 MPKTLKIRNGDKVQPTFSAGQFYARHARLRAVMAEODIEAAIFTSYHNVTYSDFLYCSF 62
QY 65 GRKYGVIYDHNNATTISAGIDGGQPMWRSSFG-DNITYTDMRRDNFRAVROLTTGAKRIG 123
DB 63 GRPYALVITQDKVVISANIDGGQPMWRSSFG-DNITYTDMRRDNFRAVROLTTGAKRIG 122
QY 124 IEPDVLDFRQLEALPGVEFVDISQPSMMKRTIKSLBEQKLIBGARVCDVGAACA 183
DB 123 VEYDHLNLQNHKGLAACPKAEIVDIAACMRKRMKSAEHHAIIRQGARVADIGAAV 182
QY 184 AATAGYPSHEEVAIATTNMIREIAKSPFVELMDTWTFQSGINTDGAHPNTNRIYOS 243
DB 183 EALREQVDEYEVAAHTQAMVREIARTYPSSELDMDTWTFQSGINTDGAHPNTNRIYOS 242
QY 244 GDIISLTFPMIFGYTALERTLFCDHVDASLDIWEKNVAVHRGELIKPGARKDIAI 303
DB 243 GEILSLNCFPMIFGYTALERTLFCDHVDASLDIWEKNVAVHRGELIKPGARKDIAI 302
QY 304 IELNEMTYREMDLTKRSFGYGSFGVLCHYRGREAGVELREDIDTBLKPGKVTVM 363
DB 303 RELNEIFLRHDLIQRFTFGYGSFGVLCHYRGREAGVELREDIDTBLKPGKVTVM 362
QY 364 LPBGPAGAGYREHDLIIYVGEAGENITGPPGGEHNIR 403
DB 363 LPBGPAGAGYREHDLIIYVGEAGENITGPPGGEHNIR 402

```

```

RESULT 6
CREA_FLASU STANDARD; PRT; 403 AA.
ID CREA_FLASU
AC P19213;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Creatinase (EC 3.5.3.3) (Creatine amidinohydrolase).
OS Flavobacterium sp. (strain U-188).
OC Bacteria; Bacteroidetes; Flavobacteriales;
OC Flavobacteriaceae; Flavobacterium.
NCBI_TaxId=242;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-15.
RX MEDLINE=91103958; PubMed=1368564;
RA Koyama Y., Kltao S., Yamamoto-Otake H., Suzuki M., Nakano E.;
RT "Cloning and expression of the creatinase gene from Flavobacterium sp.
RT U-188 in Escherichia coli.";
RL Agric. Biol. Chem. 54:1453-1457 (1990).
RN [2]
RP SEQUENCE FROM N.A.
RA Suzuki K., Sagai H., Sugiyama M., Imamura S.;
RT "Molecular cloning and high expression of the Bacillus creatinase gene
RT in Escherichia coli.";
RL J. Ferment. Bioeng. 76:77-81 (1993).
CC -1- CATALYTIC ACTIVITY: Creatine + H(2)O = sarcosine + urea.
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- SIMILARITY: To other bacterial creatinases. Also to peptidases of
CC superfamily M24.
CC -1- CAUTION: Ref.1 sequence differs from that shown due to a
CC frameshift in position 336.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D00656; BAA3239.1; -.
DR EMBL: D14464; BAA03359.1; -.
DR EMBL: D14464; BAA03360.1; ALT_TERM.
DR PIR: JH0134; JH0134.
DR HSSP: P38488; 1CHM.
DR MEROPS: M24.UNB; -.
DR InterPro: IPR000587; Creatinase.
DR InterPro: IPR000994; Peptidase_M24.
DR Pfam: PF01321; Creatinase_N; 1.
DR Pfam: PF00557; Peptidase_M24; 1.
DR KX Direct protein sequencing; Hydrolase.
FT ACT SITE 232 232 By similarity.
FT ACT SITE 232 232 By similarity.
SQ SEQUENCE 403 AA; 45780 MM; 754BDCFA56D962C CRC64;

Query Match 61.6%; Score 1349.5; DB 1; Length 403;
Best Local Similarity 64.2%; Pred. No. 2.1e-100;
Matches 257; Conservative 46; Mismatches 96; Indels 1; Gaps 1;

```

```

DB 183 EALRDVPEYEVVALHATQAMVRAIAETFDNVELMDTWTWFOSGINTDGAHPVTRKVK 242
QY 244 GDILSTWTPMIRGYTALARTLFCQHVDDASLDIWEKNVAVHRGELIKPGARCOIA 303
DB 243 GDILSTWTPMIRGYTALARTLFCQHVDDASLDIWEKNVAVHRGELIKPGARCOIA 302
QY 304 IELNEMREMDLLKYRSGYGVVCHYRGAGVLEDEDITELPKGVVSMEMPMV 363
DB 303 KELNEIRLKNDLQYRTFGVGHSPGTLSTHYGRAGLEMDIDTVLPGVVSMEMPM 362
QY 364 LPFGMPAGGYREHDLIVGEDGANTITGPPFGEHNIIR 403
DB 363 LPFGMPAGGYREHDLIVNENGAENITKFPYGERNIIR 402

RESULT 7
CREA_BACBO STANDARD; PRT; 410 AA.
ID CREA_BACBO
AC P38487;
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Creatinase (EC 3.5.3.3) (Creatine amidinohydrolase).
OS Bacillus sp. (strain B-0618).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
NCBI_TaxId=69000;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-20.
RA Suzuki K., Sagai H., Sugiyama M., Imamura S.;
RT "Molecular cloning and high expression of the Bacillus creatinase gene
RT in Escherichia coli.";
RL J. Ferment. Bioeng. 76:77-81 (1993).
CC -1- CATALYTIC ACTIVITY: Creatine + H(2)O = sarcosine + urea.
CC -1- SUBUNIT: Homodimer.
CC -1- INDUCTION: By choline chloride.
CC -1- SIMILARITY: To other bacterial creatinases. Also to peptidases of
CC superfamily M24.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D14463; BAA03358.1; -.
DR PIR: I39809; I39809.
DR HSSP: P38488; 1CHM.
DR InterPro: IPR000587; Creatinase.
DR InterPro: IPR000994; Peptidase_M24.
DR Pfam: PF01321; Creatinase_N; 1.
DR Pfam: PF00557; Peptidase_M24; 1.
DR KX Direct protein sequencing; Hydrolase.
FT INIT MET 0
FT ACT SITE 239 0 By similarity.
FT ACT SITE 239 0 By similarity.
SQ SEQUENCE 410 AA; 46618 MM; AAA2BFC1067B1A22 CRC64;

Query Match 61.6%; Score 1349; DB 1; Length 410;
Best Local Similarity 63.9%; Pred. No. 2.4e-100;
Matches 253; Conservative 49; Mismatches 90; Indels 4; Gaps 1;

```


QY 188 AGVPEHEVALATNAMI REIAKSPFVPELMDTWTFOSGINTDGAHNPVTNRIVOSGDI 247
 DB 194 EGVPEHEVALHGEAMVREIARTYPAELRDTIMVQSGINTGAINMATSRGQGDIL 253
 QY 248 SLNTPFMIFGYTALERTLFCDHVDASLDIMKRNVAHRRGLELIKPGARCKDIAIEIN 307
 DB 254 SLNCFPMIAGYTALERTLFLIEVSDRHELMWINKVRHGLELIKPGARCKDIAIEIN 313
 QY 308 EAVREMDLKYRSFGYSHFGVLCYHRYGAGVIELDEDIDELKPGMVMSEMMVLPFG 367
 DB 314 EIVREHDLANRPFYGHSGFVLSHYGRGAGLELEDIETVLEPGMVMSEMMVLPFG 373
 QY 368 MPAGGGRHDLIVGEDGAEINTGPPFGEHNIR 403
 DB 374 EPGAGGRHDLIVISENGENTTKPFGEHNIR 409

RESULT 8
 CREA_PSEPU STANDARD; PRT; 403 AA.
 AC P38488;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Creatinase (EC 3.5.3.3) (Creatine amidinohydrolase).
 OS Pseudomonas putida.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=303;
 RN [1]
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
 RX MEDLINE=90339496; PubMed=1696320;
 RA Hoeftken H.W., Knof S.H., Bartlett P.A., Huber R., Moellerling H.,
 "Enzymatic mechanism of creatine amidinohydrolase as deduced from
 crystal structures.";
 RT J. Mol. Biol. 214:597-610 (1990).
 RL [2]
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
 RX MEDLINE=89125596; PubMed=3221393;
 RA Hoeftken H.W., Knof S.H., Bartlett P.A., Huber R., Moellerling H.,
 "Crystal structure determination, refinement and molecular model of
 creatine amidinohydrolase from Pseudomonas putida.";
 RT J. Mol. Biol. 204:417-433 (1988).
 RL [1]
 CC -1 CATALYTIC ACTIVITY: Creatine + H(2)O = sarcosine + urea.
 CC -1 SUBUNIT: Homodimer.
 CC -1 DOMAIN: Each monomer has two clearly defined domains. The small N-
 terminal domain (AA 1-161) and the large domain (AA 162-403). Each
 of the two active sites is made by residues of the large domain of
 one monomer and some residues of the small domain of the other
 monomer.
 CC -1 SIMILARITY: To other bacterial creatinases. Also to peptidases of
 superfamily M24.
 CC PDB: 1CHM; X-Ray; A/B=2-402.
 DR InterPro: IPR000587; Creatinase.
 DR InterPro: IPR000994; Peptidase_M24.
 DR Pfam: PF01321; Creatinase_N; 1.
 DR Pfam: PF00557; Peptidase_M24; 1.
 KW 3D-structure; Hydrolase.
 FT ACT SITE 232 232
 FT STRAND 6 7
 FT HELIX 21 37
 FT TURN 38 39
 FT STRAND 42 45
 FT HELIX 48 55
 FT TURN 62 63
 FT STRAND 67 70
 FT STRAND 75 80
 FT HELIX 81 83
 FT TURN 84 85
 FT HELIX 86 90

FT STRAND 95 99
 FT TURN 101 102
 FT TURN 104 105
 FT HELIX 106 114
 FT STRAND 120 123
 FT TURN 125 127
 FT STRAND 129 129
 FT HELIX 130 139
 FT TURN 141 142
 FT STRAND 144 147
 FT HELIX 149 156
 FT TURN 157 157
 FT HELIX 161 184
 FT TURN 187 188
 FT STRAND 190 190
 FT STRAND 191 209
 FT STRAND 215 215
 FT STRAND 219 224
 FT HELIX 225 229
 FT TURN 231 232
 FT STRAND 235 236
 FT STRAND 239 239
 FT TURN 242 243
 FT STRAND 245 251
 FT STRAND 253 254
 FT TURN 255 256
 FT STRAND 257 258
 FT STRAND 261 267
 FT STRAND 272 291
 FT HELIX 294 295
 FT TURN 297 297
 FT STRAND 298 312
 FT TURN 313 313
 FT HELIX 315 317
 FT STRAND 318 318
 FT STRAND 324 325
 FT STRAND 328 330
 FT TURN 331 332
 FT STRAND 333 334
 FT TURN 337 338
 FT STRAND 340 340
 FT TURN 343 344
 FT STRAND 348 348
 FT TURN 351 352
 FT STRAND 354 357
 FT TURN 360 363
 FT STRAND 365 366
 FT TURN 368 369
 FT STRAND 371 374
 FT STRAND 376 382
 FT TURN 383 384
 FT STRAND 385 388
 FT TURN 396 399
 FT HELIX 400 401
 SQ SEQUENCE 403 AA; 45536 MM; 21C2CD080BD017E0 CRC64;

Query Match 61.5%; Score 1347.5; DB 1; Length 403;
 Best Local Similarity 64.0%; Pred. No. 3e-100;
 Matches 256; Conservative 47; Mismatches 96; Indels 1; Gaps 1;

QY 5 MLHWKMGNGEKDPSFSDAEMTRONDVRGMKANVNDALFTSYHCINTYSGWLYCF 64
 DB 3 MKTIRINGDQVNSTFSAQEIYANKQARLRAHIAENIDALFTSYHNINYSDFLYCSF 62
 QY 65 GRKYGMVIDHNNATTISAGIDGQFWRSSFG-DNITYTDMRDNFRAVRQLTGAKIG 123
 DB 63 GRPYALVTBEDVVISANIDGGQPMWRTVTGDNIVYTDWQRDNVFAALQALPKARRIG 122
 QY 124 IEPDVNIDFPRQLEBALPGVEFVDISQPSMMWRTTKLEQKLRGQAVCDVGAACA 183
 DB 123 IEHDLNLQNRDKLAARYPDALVDVAACRMRMVKSABEHVMIIRHGRLADIGGAIVV 182
 QY 184 AAKAGVEHEVALATNAMI REIAKSPFVPELMDTWTFOSGINTDGAHNPVTNRIVOS 243

DB 183 EALGQVEYEVYALATQAWRAIADTFEDVEHMTWTWTFQSGINTIDGAAHVPYTRKXK 242
QY 244 GDILSLNTPFPMIFGYTALERTLFCDDHYDASLDIWEKNVAHRRGLBLIKPGARCKDIA 303
DB 243 GDILSLNTPFPMIAGYTTALERTLFLDHCSDHLRLMQVVEHAGLKLKPGARCKDIA 302
QY 304 IELNMTREMDLKTFRSGYGHSPVLCHYYREKGVLEIRIDTELKPGMWVSNPPVM 363
DB 303 RELNDEIFLKHVDYLQYRTFGYGHSPGTLSHYYGREAGLEIRIDITVLEPGMWVSNPPVM 362
QY 364 LPEGMGAGYREHDIILVGEDGAEINITGFPFGEHNIIR 403
DB 363 LPEGLPGAGYREHDIILVNGEAEINITGFPGEHNIIR 402

RESULT 9

066182 PRELIMINARY; PRT; 411 AA.
AC 066182;
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Creatinase.
OC Archaeobacter sp.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Micrococcineae; Micrococcaceae; Archaeobacter.
OX NCBI_TaxID=1667;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TE1826;
RA Nishiyama Y., Imanaka T.;
RT "Analysis of a negative regulator, soxR, for the Archaeobacter
sarcosine oxidase gene."
RL J. Ferment. Bioeng. 81:64-67(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=TE1826;
RX MEDLINE=95223334; PubMed=9563845; DOI=10.1007/s004380050685;
RA Nishiyama Y., Toda A., Imanaka T.;
RT "Gene cluster for creatinine degradation in Archaeobacter sp. TE1826."
RL Mol. Genet. 257:581-586(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=TE1826;
RA Nishiyama Y., Imanaka T.;
RT "Cloning and sequencing of the sarcosine oxidase gene from
Archaeobacter sp. TE1826."
RL J. Ferment. Bioeng. 75:239-244(1993).
DB BMBL; AB007122; BAA25928.1; --
DR HSP; P38488; ICHM.
DR GO; GO:0016980; F:creatinase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0008235; F:metalloexopeptidase activity; IEA.
DR GO; GO:0006600; P:creatinase metabolism; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000587; Creatinase.
DR InterPro; IPR000994; Peptidase M24.
DR Pfam; PF01321; Creatinase N; 1.
DR Pfam; PF00557; Peptidase M24; 1.
KW Hydrolase.
SQ SEQUENCE 411 AA; 47147 MW; B8C5DS9A886DE201 CRC64;

Query Match 60.7%; Score 1329.5; DB 2; Length 411;
Best Local Similarity 63.0%; Pred. No. 8.9e-99;
Matches 250; Conservative 55; Mismatches 87; Indels 5; Gaps 3;

QY 12 HNGEKOYSPSPDAENTRRQNDVGRGMKKNVDAALFTSYHCINYYSGMLYCTFGKRYGV 71
DB 14 HNGEKFQGFSPKXENDRRNTNLRNVAEKDIDAVLFTSYHNINYYSDFLYTSFNNRYGLV 73
QY 72 IDHNAATISAGIDGQPPWRSGDNITVTWRDRDFFYAVRQL--TTGAK--RIGIED 127

DB 74 VTQKHTVTSANIDGKPPWRSDYDENITVYTDNRDNYFALQKLEAGVKKARLIGIED 133
QY 128 HVNIDFRQLEALPGVEFVDISQPSMMNRITKLEQKILREGARVCDVGGAACAAAIK 187
DB 134 HVSIDLKRFSDTFPNRELVAHVSDVMKQRIKSAEELRHKNGARLINDIGYAVVAIQ 193
QY 188 AGVEPEHVALATYTNAMIREIKSPFVELMDTWTWTFQSGINTDGAHNPVTRIVOSGDIL 247
DB 194 EGVEPEYVALAGSOMTRERIKALYPOSELRTWTWTFQSGINTDGAHNSWATSKVQKSEL 253
QY 248 SLNTPFPMIFGYTALERTLFCDDHYDASLDIWEKNVAHRRGLBLIKPGARCKDIAELN 307
DB 254 SLNTPFPMIAGYTTALERTLFLAEVSDAHLKYWEINVEYHKGLELTKPGAVCKDIAELN 313
QY 308 EMYREMDLKTFRSGYGHSPVLCHYYREKGVLEIRIDTELKPGMWVSNPPVMLEPG 367
DB 314 EMPREHDLVKNRTFGYGHSPGVLSHYYGREAGLEIRIDITVLEPGMWVSNPPVMLEPG 373
QY 368 MEGAGYREHDIILVGEDGAEINITGFPFGEHNIIR 403
DB 374 MEGAGYREHDIILVQENGVEDITGFPFGEHNIIR 410

RESULT 10

086ML3 PRELIMINARY; PRT; 376 AA.

ID 086ML3
AC 086ML3;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Cobalt dependent X-pro dipeptidase (EC 3.4.13.9).
CN OrderedLocustNames=OB2829;
OS Oceanobacillus thelyensis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.
OX NCBI_TaxID=182710;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HTE831;
RX MEDLINE=22220767; PubMed=12235376; DOI=10.1093/nar/gkf526;
RA Takami H., Takaki Y., Uchiyama I.;
RT "Genome sequence of Oceanobacillus thelyensis isolated from the Iheya
Ridge and its unexpected adaptive capabilities to extreme
environments."
RL Nucleic Acids Res. 30:3927-3935(2002).
DR EMBL; AP004602; BAC14785.1; --
DR GO; GO:0004251; F:X-pro dipeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000994; Peptidase M24.
DR Pfam; PF00557; Peptidase M24; 1.
KW Complete proteome; Hydrolase.
SQ SEQUENCE 376 AA; 42679 MW; 046EAB64D13E0AD2 CRC64;

Query Match 12.8%; Score 279.5; DB 2; Length 376;
Best Local Similarity 23.4%; Pred. No. 4.3e-14;
Matches 92; Conservative 74; Mismatches 183; Indels 45; Gaps 10;

QY 29 RQNDVGRGMKKNVDAALFTSYHCINYYSGMLYCTFGKRYGVIDHNAATISAGIDGQ 88
DB 6 RVDNRLALMERSNIDLSIMNVENOYINGKATYSPYVALDSKSLSLTPSLEENH 65
QY 89 PWRSGDNITVYTDNRDNYFALQKLEAGVKKARLIGIEDHVNLD 134
DB 66 AKETIDANELYIYHETNLYGDKSYDY----LETYSQYPNGT-RVGVEPSSLSLMA 120
QY 135 RQLEBALPGVEFVDISQPSMMNRITKLEQKILREGARVCDVGGAACAAAIKAGVPEHE 194
DB 121 NILKKA--GPELINLDHRIAEERIKYONDEEILEMVOEAKGLSLAKSLLENVAQPCITWE 178
QY 195 VALATYTNAMIREIKSPFVELMDTWTWTFQSGINT-----DGAHNPVTRIVOSGDIL 247
DB 179 LDRHGTQLLEFEM--DNYPDATL-DYFANSPGSLERSIMPHVFSNTRKLLINDIVHSHOV 236
QY 248 SLNTPFPMIFGYTALERTLFCDDHYDASLDIWEKNVAHRRGLBLIKPGARCKDIAELN 307

Db 227 ALN-----GYRACERTFFVGKPTDKQDAFAKAAVEAQLVAMEKJAVGTAKEDVAR 290
QY 308 EHYREMDLTKRSFGYGSFGVLCHYGRBAGVELREDIDTLKSGMVSMPPMMLPBG 367
Db 291 NIFQGYNEKCIHRTGHGIGI-----GIHEPSRLFNDVLQSGMFTIEPGIYI--- 342
QY 368 MPGAGYREHDLIVGEDGAENITGPPGPEHNI 401
Db 343 -PGVGGRHSDTVVLTNEGTLVTEYPREIEHLI 375

RESULT 11
06SKC1 PRELIMINARY; PRT; 364 AA.
AC 06SKC1; 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE Ykwy (Peptidase M24).
GN Name=Ykwy; ORFNames=BL03561, BL01594,
OS Bacillus licheniformis DSM 13.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=279010;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 13.
RX PubMed=15383718;
RA Velth B., Herzberg C., Steckel S., Feesche J., Maurer K.H.,
RA Ehrenreich P., Baumeier S., Henne A., Liesegang H., Merkl R.,
RA Ehrenreich A., Gottschalk G.;
RT "The Complete Genome Sequence of Bacillus licheniformis DSM13, an
RT Organism with Great Industrial Potential.",
RL J. Mol. Microbiol. Biotechnol. 7:204-211(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 14580;
RA Rev M.W., Ranaiva P., Nelson B.A., Brody-Karpin S.D., Zaretsky E.J.,
RA Tang M., de Leon A.L., Xiang H., Gusti V., Clausen I.G., Olsen P.B.,
RA Raasmussen M.D., Andersen U.T., Jorgensen P., Larsen T.S.,
RA Sorokin A., Bolotin A., Lapidus A., Galleron N., Ehrlich S.D.,
RA Berka R.M.;
RT "Complete genome sequence of the industrial bacterium Bacillus
RT licheniformis and comparisons with closely related Bacillus species.",
RL Genome Biol. 5:R77-R77(2004).
DR EMBL; AB017333; AA040493.1; -
DR EMBL; CP000002; AAU23135.1; -
SQ SEQUENCE 364 AA; 40396 MW; 2ADD117D8D34B1C7 CRC64;

Query Match 12.5%; Score 274.5; DB 2; Length 364;
Best Local Similarity 23.3%; Pred. No. 11e-13;
Matches 94; Conservative 63; Mismatches 161; Indels 85; Gaps 11;

QY 29 RQNDVRGMMAKNVDAALFTS---YHCINYY----- 56
Db 3 RIEQISSLSLKKRNTSSAFHSEKENVFYLTGTEPHERLMGVITFQDEPPFVCGQWAG 62
QY 57 ----SGWLICYFGKRYGVNIDHNNATTSAGIDGQPPRRSGDNITTTWRDNFYRAV 112
Db 63 QARAAGAYEIIIG--YG---DHQN-----PM-----ELISSAL 90
QY 113 RQLTTGAARIGIEPDHVVLDPRROLEALPGVEPVDISQPSMMRTIKSLBEQKLIKREGA 172
Db 91 QKRNGQCLKVAVEKTLTASRAEMQAATNGNELVSABERLNQIRVTDKKEISILREAA 150
QY 173 RVCDVGGAAACAATKAGVPEHEVAIATNNAMIR--EIAKSPFVELMDTWTWFOGSGINTDG 231
Db 151 KLAAYGVRAAGAAALKEGIAIDV-----AKIEYTLKKK--GVQMSSTVTLRGKSGQ 203
QY 232 AHNPTVNRIVOSGDILSINTEPMIFGYTALERTLFCDHVDASLDIWEKNVAVHRRGLE 291
Db 204 PHGNPGRSTLRKPGDFVLPDLGVITIDGYCSDIRTLVYQNVSEKQEIYNTVTLQATEALK 263

QY 292 LKPGARCKDIAIEINEMREMDLTKRSFGYGSFGVLCHYGRBAGVELREDIDTLK 351
Db 264 MSKPGVRIGDDLDKARGLIEKAGYGDYFPNHLGHGLGSPHEX-----PSMHNDELK 318
QY 352 PGMVSMPPMMLPBGMPGAGYREHDLIVGEDGAENITGPP 394
Db 319 QGMVYTIIEPGIYIPE-----IGVRIEDDVLVTADGAAALTKVP 357

RESULT 12
067N93 PRELIMINARY; PRT; 357 AA.
AC 067N93; 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE Xaa-Pro dipeptidase.
GN ORFNames=STH1865;
OS Symbiodacterium thermophilum.
OC Bacteria; Actinobacteria; Symbiodacterium.
OX NCBI_TaxID=2734;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IAM14863;
RA Ueda K., Yamashita A., Ishikawa J., Shinada M., Watsuji T.,
RA Morimura K., Ikeda H., Hattori M., Beppu T.;
RT "Complete genome sequence of an uncultured bacterium Symbiodacterium
RT thermophilum.",
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP006840; BAD04850.1; -
DR InterPro; IPR000994; Peptidase M24.
DR InterPro; IPR001714; Pept M24 MAP.
DR Pfam; PF00557; Peptidase M24; 1.
DR PRINTS; PRO0599; MAPEPTIDASE.
KW Hydrolase.
SQ SEQUENCE 357 AA; 39193 MW; 43A1576729876460 CRC64;

Query Match 12.5%; Score 273.5; DB 2; Length 357;
Best Local Similarity 25.1%; Pred. No. 1.2e-13;
Matches 97; Conservative 65; Mismatches 175; Indels 49; Gaps 10;

QY 22 SDAEMTRQNDVRGMMAKNVDAALFTSYHCINYSGM-----LYCYFGKRYGVNIDHNN 76
Db 2 SQARLAR-----LRARLERGLDAVLAKPHRAVLSTGTSAGLLITADAAALVTDRY 57
QY 77 ATTSAGIDGQPPRRSGDNITTT---DWRDNFYRAVQLTTGAKRIGIEPDHVLND 132
Db 58 VEQAAQAQPEPEVIAKASANOAVLTRVVEW-----GVKRIEFGDFLTVD 103
QY 133 PRROLEALPGVEPVDISQPSMMRTIKSLBEQKLIKREGAVCDVGGAAACAATKAGVPE 192
Db 104 EHQYQQLGACQWTSVSGVLEBELRMKIDETELMRMAATADBAFAQIILPLKPGVIE 163
QY 193 HEVAIATNNAMIRIASSFPFVELMDTWTWFOGSGINTDGANPTVNRIVOSGDILSLNTF 252
Db 164 RDLATELEYNRKKLGAGVAFETIV-----ASGARSPLPGVASDKAIEVGDITLTPFG 217
QY 253 PMIFGYTALERTLFCDHVDASLDIWEKNVAVHRRGLEIKPGARCKDIAIEINEMRE 312
Db 218 AVYQGYCSMTFTWMLGEPTDKREIYGIVLEAKRGVAACRPGITGR-----ELDDVRS 273
QY 313 MDLKL-YRSF---GYGHSFGVLCHYGRBAGVELREDIDTLKSGMVSMPPMMLPBG 368
Db 274 YIAKGYREYFGHGTGAGVYIHEGPRVS---QRGDVLVLRGMVTVTEPGIYL----- 325
QY 369 PGAGYREHDLIVGEDGAENITGPP 394
Db 326 PGMGVRIEDMLVTEBSGASFTHSP 351

RESULT 13
065HH3 PRELIMINARY; PRT; 353 AA.

AC 065H3;
 DT 25-OCT-2004 (TReMBLrel. 28, Created)
 DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
 DE Yght (Peptidase M24B, X-Pro dipeptidase).
 GN Name=yght; ORFNames=BL01543, BL102617;
 OS Bacillus licheniformis DSM 13
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxId=279010;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DSM 13;
 RX PubMed=15383718;
 RA Veith B., Herzberg C., Steckel S., Feesche J., Maurer K.H.,
 RA Ehrenreich P., Baumeier S., Henne A., Liesegang H., Weikl R.,
 RA Ehrenreich A., Gottschalk G.,
 RA "The Complete Genome Sequence of Bacillus licheniformis DSM13, an
 RT Organism with Great Industrial Potential."
 RL J. Mol. Microbiol. Biotechnol. 7:204-211(2004).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 14580;
 RA Rev M.W., Ramaïya P., Nelson B.A., Brody-Karpin S.D., Zaretsky E.J.,
 RA Tang M., de Leon A.L., Xiang H., Guest V., Clausen I.G., Olsen P.B.,
 RA Rasmussen M.D., Andersen J.T., Jorgensen P.L., Larsen T.S.,
 RA Sorkin A., Bolotin A., Lapidus A., Galleron N., Ehrlich S.D.,
 RA Berka R.M.;
 RA "Complete genome sequence of the industrial bacterium Bacillus
 RT licheniformis and comparisons with closely related Bacillus species";
 RL Genome Biol. 5:R77-R77(2004).
 DR EMBL; AB017333; AAU41491.1;
 DR EMBL; CP000002; AAU24132.1;
 DR EMBL; CP000002; AAU24132.1;
 SQ SEQUENCE 353 AA; 38339 MW; 2EC5DCED99160647 CRC64;
 Query Match 12.3%; Score 270; DB 2; Length 353;
 Best Local Similarity 25.8%; Pred. No. 2,3e-13;
 Matches 99; Conservative 58; Mismatches 162; Indels 64; Gaps 11;

QY 31 NDVGMKANNVDAALFTSYHCINYSGLYCYFGRKYGWIDHNNATTISAGIDGGOPW 90
 DB 4 NKEHLEFGGLGIDGLVTSYGVNLRYITG---FTGSSGLAVISDQAVFIR----- 50
 QY 91 RRSRGDNTTYDWRNDNFYR-----AVROLTY-----GAKRIGIEPDVNLDFRRQ 136
 DB 51 -----DRTTYEQADQIKSFYIOHAGIYQKTAIVIKEMGINKIGFEDKRYMTYTAA 104
 QY 137 LEEALPGVEFYDISOPSMWRITKSLSEOKLIREGARVCVGGAAACAARAGVPEHEVA 196
 DB 105 YKEQGLAALVYPVSGSVKRLIKSSSEIKILKEAKIADAFSHILVILVIGITIEIVA 164
 QY 197 IATTNAMIETIA--KSPFVELMDTWTFQSGINTDGAHNPVTNRIVOSGILSLNTPFM 254
 DB 165 NELFMYRSGADHSSFDMI-----VASGVNLSLPHGVAAGVAKIEKGLVTLDFPAY 216
 QY 255 IFGYTALERTLFCCHVDASLDIWEKVAVHRRLGLIKPGARKDIAILENMYREMD 314
 DB 217 YKGYCSDTITRYVAEGPDELKRIYQTVFEAALGWSIKPGITCK---PADATRY- 271
 QY 315 LLKTRSFYGHSPVLCHYRGREAGVELRED-----IDTELPGMVVSMPEMWLPEGM 368
 DB 272 ---ISSQYGVYFG---HSTGHGLGMEVHESBALSARDQMLEKGMVTVTFEGIYI--- 321
 QY 369 PGAGGYRHDILVGEDGAENIT 391
 DB 322 PGKGVRIEDDILVTEGNEISLT 344

RESULT 14
 Q722CO PRELIMINARY; PRT; 365 AA.
 ID Q722CO;
 DT 05-JUL-2004 (TReMBLrel. 27, Created)
 DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)

DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
 DE Proline dipeptidase (EC 3.4.13.9).
 GN Name=pepo; OrderedlocusNames=BCE4748;
 OS Bacillus cereus (Strain ATCC 10987).
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxId=222523;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC PubMed=1960714; DOI=10.1093/nar/gkh258;
 RA Raeko D.A., Ravel J., Oestad O.A., Helgason E., Cer R.Z., Jiang L.,
 RA Shores K.A., Fouts D.E., Tourasse N.J., Anguilo S.V., Kolonay J.F.,
 RA Nelson W.C., Kjolseth A.-B., Frazer C.M., Read T.D.;
 RA "The genome sequence of Bacillus cereus ATCC 10987 reveals metabolic
 RT adaptations and a large plasmid related to Bacillus anthracis pXO1";
 RL Nucleic Acids Res. 32:977-988(2004).
 DR EMBL; AB017279; AAS43649.1;
 DR HSSP; BCS4748;
 DR TIGR; P15034; 1A16.
 DR GO; GO:0004239; F:methionyl aminopeptidase activity; IEA.
 DR GO; GO:0004251; F:X-Pro dipeptidase activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR000994; Peptidase_M24.
 DR InterPro; IPR001131; Peptidase_M24B.
 DR InterPro; IPR001714; Pept_M24_MAP.
 DR Pfam; PF00557; Peptidase_M24; 1.
 DR PRINTS; PR00599; MAPPEPTIDASE.
 DR PROSITE; PS00491; PROLINE_PEPITIDASE; 1.
 DR Complete proteome; HydroLase.
 KW SEQUENCE 365 AA; 40636 MW; 9EC4FBC713B9BED CRC64;
 Query Match 11.9%; Score 261.5; DB 2; Length 365;
 Best Local Similarity 25.1%; Pred. No. 1,2e-12;
 Matches 95; Conservative 67; Mismatches 187; Indels 29; Gaps 12;

QY 26 MTRRQNDVRCGMANNVDAALFTSYHCINYSGLYCYFGRKYGWIDHNNATTISAG 83
 DB 1 MNARLEIMQMLKEKNEVAALFTSTPVFYMTN-ICEPHERLIGMVFQEKELILCPK 59
 QY 84 IDGGPFRBSPGDNIT---TTDWRDNPFRVAVQLTGARRIGIEPDVNLDFRRQLEA 140
 DB 60 MEEQANAGMAHEITIGFTIDRPWDITATIKDGINNANAVALEKHELVNEREELTKL 119
 QY 141 LPGEFYDISOPSMWRITKSLSEOKLIREGARVCVGGAAACAARAGVPEHEVAIAT 200
 DB 120 FPNAAFTSAEKRELRILDEKELSLREAAQADVAVEGVNAIKEDSELEV----- 175
 QY 201 NAMIREIAKSPFVELMDTWTFQSGINTDGAHNPVTNRIVOSGILSLNTPFMITGY 259
 DB 176 -AIEHEHLK-EGIHKMSFTMTVLGANSALPHGIPGANK-MKRGDFVLFDLGVITDGYC 232
 QY 260 TALERTLFCCHVDASLDIWEKVAVHRRLGLIKPGARKDIAILENMYRE---WDL 316
 DB 233 SDITRYAFDISIEQTRINTVLAQGLQAVEACKPVTTLGALDINAARSVIADGYGDF 292
 QY 317 KYRSGFYGHSPVLCHYRGREAGVELREDIDTELPGMVVSMPEMWLPEGMGAGYRE 376
 DB 293 PHR---LGHGLGISVHY-----PDVKAQNSPLKEGMVTVTFEGIYI---VPVGVRI 340
 QY 377 HDILVGEDGAENITGFP 394
 DB 341 EDDIYITKDSSEILTKFP 358

RESULT 15
 Q6HCR7 PRELIMINARY; PRT; 365 AA.
 ID Q6HCR7;
 DT 05-JUL-2004 (TReMBLrel. 27, Created)
 DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
 DE Xaa-Pro dipeptidase (Proline dipeptidase) (EC 3.4.13.9).
 GN Name=pepo; OrderedlocusNames=BT9727_4344;
 OS Bacillus thuringiensis (subsp. konkukian).

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=180856;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-97-27;
RA Brettn T.S., Bruce D., Challacombe J.F., Gilna P., Han C., Hill K.,
RA Hitchcock P., Jackson P., Kelm P., Longmire J., Lucas S., Okinaka R.,
RA Richardson P., Rubin E., Rice H.,
RT "Complete genome sequence of *Bacillus thuringiensis* 97-27.";
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL, AE017355; AA163552.1; ..
DR GO; GO:0004239; F:methionyl aminopeptidase activity; IEA.
DR GO; GO:0004251; F:X-Pro dipeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000994; Peptidase_M24.
DR InterPro; IPR001131; Peptidase_M24B.
DR InterPro; IPR001714; Pept_M24_MAP.
DR Pfam; PF00557; Peptidase_M24; 1.
DR PRINTS; PR00599; MAPEPTIDASE.
DR PROSITE; PS00491; PROLINE PEPTIDASE; 1.
KW Complete proteome; Hydrolase.
SQ SEQUENCE 365 AA; 40668 MW; 36A6B2971F3E5CE CRC64;

Query Match 11.9%; Score 261.5; DB 2; Length 365;
Best Local Similarity 24.9%; Pred. No. 1.2e-12;
Matches 94; Conservative 65; Mismatches 191; Indels 27; Gaps 10;

QY 26 MTRQNDVRGMMAKNNVDAALFTSYHCINYSGLICYFGKRYGM-VIDHNNATTISAGI 84
DB 1 MNAIRENLMQMLKENQVEAAFLSTPNVFYMTNFCPEHERLGMFVQEKEPILICPKM 60
QY 85 DGGQWRRSFGDNI---TYTDWRDNFYRAVAVQLTTGAKRIGIEFDHYNLDFRQLEBAL 141
DB 61 EEGQARNAGMAHEITIGFTDTRPMDIAKAIKDRGINANAVAIKEMHINVERYEELTKLP 120
QY 142 PGVEFVDISQPSMMWRITIKSLSEQQLIREGARVCDVGAAACAALIKAGVPEHEVAIATTN 201
DB 121 PNAAFTSAEKVRLELRIKDEKELSLREAAAMADYAVEGVNAIKENRSELEVL----- 175
QY 202 AMIREIASPPFVELMDTWTFQSGINTDGAHN-PVTRIVOSGILSLNTPPMIFGYTT 260
DB 176 AIIHEHLEKT-KGIHMKSPDTWVLGANSALPHGIPGANK-MNRGDFVLFDLGVIIDGYCS 233
QY 261 ALERTLFCDHVDDASLDIMEKNVAVHRGRLRIKPGARCKDAIELNEMVRE---WDLK 317
DB 234 DITRTVAFGEISSEQTRITNTVTLAQLOAVEACKPGVTLGALDNNARSVIADAGYGFPP 293
QY 318 YRSFGYGHSPFVLCYVYGREAGVELREDIDTELKPGMVSMEPVVMLPEGMPGAGYREH 377
DB 294 HR---LGHGLGISVNEY-----PDVKAGNESPLKKGWFTTEPGIY---VPNVGVRIE 341
QY 378 DILIVGEDGAENTIGFP 394
DB 342 DDIVYITKDGSEILTKFP 358

Search completed: July 8, 2005, 01:35:42
Job time : 87 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 8, 2005, 01:18:01 ; Search time 27 Seconds
(without alignments)
1116.971 Million cell updates/sec

Title: US-10-807-228a-1

Perfect score: 2190

Sequence: 1 MTDMDLHVKNMNGEKDYSY.....DGAENITGPFGEHNIIRN 404

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 513545 seqs, 7464964 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/prodata/1/aa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/aa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/aa/6A_COMB.pep:*
4: /cgn2_6/prodata/1/aa/6B_COMB.pep:*
5: /cgn2_6/prodata/1/aa/PCTUS_COMB.pep:*
6: /cgn2_6/prodata/1/aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysts of the total score distribution.

SUMMARIES

Result No.	Score	* Match	Query Length	DB ID	Description
1	2190	100.0	404	1 US-09-940-941-1	Sequence 1, Appl1
2	2190	100.0	404	2 US-08-947-726A-2	Sequence 2, Appl1
3	2190	100.0	404	3 US-08-799-897-1	Sequence 1, Appl1
4	2190	100.0	404	4 US-09-856-645-1	Sequence 1, Appl1
5	2190	98.1	404	4 US-09-869-280A-1	Sequence 1, Appl1
6	236.5	10.8	362	4 US-09-134-000C-4306	Sequence 4306, Ap
7	219	10.0	366	4 US-09-107-532A-4292	Sequence 4292, Ap
8	200.5	9.2	370	3 US-09-134-001C-3769	Sequence 3769, Ap
9	196.5	9.0	390	3 US-09-134-001C-3627	Sequence 3627, Ap
10	188.5	8.6	360	4 US-09-583-110-3285	Sequence 3285, Ap
11	188.5	8.6	361	4 US-09-107-433-3390	Sequence 3390, Ap
12	187.5	8.6	363	4 US-09-602-777A-94	Sequence 94, Appl1
13	183	8.4	353	4 US-09-583-110-4793	Sequence 4793, Ap
14	183	8.4	358	4 US-09-107-433-3157	Sequence 3157, Ap
15	181.5	8.3	299	4 US-09-861-451A-36	Sequence 36, Appl1
16	171	7.8	355	4 US-09-198-452A-871	Sequence 871, App
17	171	7.8	364	4 US-09-438-185A-815	Sequence 815, App
18	160.5	7.3	253	4 US-09-523-263B-10	Sequence 10, Appl1
19	147	6.7	318	4 US-09-107-532A-5738	Sequence 5738, Ap
20	143.5	6.6	389	4 US-09-797-906-4	Sequence 4, Appl1
21	143	6.5	379	4 US-09-602-777A-112	Sequence 112, App
22	142.5	6.5	289	3 US-09-797-906-5	Sequence 5, Appl1
23	142	6.5	274	4 US-09-523-263B-11	Sequence 11, Appl1
24	140.5	6.4	248	4 US-09-523-263B-15	Sequence 15, Appl1
25	140.5	6.4	485	4 US-09-902-540-9878	Sequence 9878, Ap
26	140	6.4	441	4 US-09-543-681A-5018	Sequence 5018, Ap
27	138	6.3	520	4 US-09-724-623-124	Sequence 124, App

28	135	6.2	416	4 US-09-252-991A-26713	Sequence 26713, A
29	134	6.1	394	4 US-09-523-263B-2	Sequence 2, Appl1
30	133	6.1	387	4 US-09-523-263B-9	Sequence 9, Appl1
31	133	6.1	492	4 US-09-489-039A-11838	Sequence 11838, A
32	128.5	5.9	260	4 US-09-902-540-12329	Sequence 12329, A
33	128	5.8	264	4 US-09-328-352-7576	Sequence 7576, Ap
34	127.5	5.8	265	4 US-09-543-681A-7847	Sequence 7847, Ap
35	127.5	5.8	441	4 US-09-492-709A-272	Sequence 272, App
36	124.5	5.7	250	4 US-09-523-263B-22	Sequence 22, Appl1
37	123.5	5.6	252	4 US-10-138-701-26	Sequence 26, Appl1
38	121.5	5.5	389	4 US-09-270-767-41600	Sequence 41600, A
39	117	5.3	305	4 US-09-523-263B-14	Sequence 14, Appl1
40	116	5.3	1452	2 US-08-652-971-4	Sequence 4, Appl1
41	116	5.3	1452	2 US-08-991-258A-4	Sequence 4, Appl1
42	116	5.3	1452	2 US-08-769-399-4	Sequence 4, Appl1
43	116	5.3	1452	3 US-08-991-953A-4	Sequence 4, Appl1
44	115.5	5.3	285	4 US-09-252-991A-20742	Sequence 20742, A
45	114.5	5.2	268	4 US-09-523-263B-13	Sequence 13, Appl1

ALIGNMENTS

RESULT 1
US-09-940-941-1
; Sequence 1, Application US/09940941
; Patent No. RE38687
; GENERAL INFORMATION:
; APPLICANT: Sogabe, Atsushi
; Hattori, Takashi
; Nishiy, Yoshiaki
; Kawamura, Yoshihisa
; TITLE OF INVENTION: NOVEL CREATINE AMIDOHYDROLASE, PRODUCTION
; THEREOF AND USE THEREOF
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LEYDIG, VOIT & MAYER, LTD.
; STREET: Two Prudential Plaza, Suite 4900
; CITY: Chicago
; STATE: Illinois
; COUNTRY: US
; ZIP: 60601-6780
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/940,941
; FILING DATE: 26-Aug-2001
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/799,897
; FILING DATE: 13-FEB-1997
; APPLICATION NUMBER: JP 25435/1996
; FILING DATE: 13-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Robert F. Green
; REGISTRATION NUMBER: 27555
; REFERENCE/DOCKET NUMBER: 78064
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 616-5700
; TELEFAX: (312) 616-5700
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 404 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: <Unknown>
; DESCRIPTION: protein
; ORIGINAL SOURCE:
; ORGANISM: Alcaligenes faecalis

STRAIN: TE3581 (FERM P-14237)
FEATURE:
NAME/KEY: mat peptide
LOCATION: 1 to 404
OTHER INFORMATION: protein having creatine amidino-
hydrolase activity
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-940-941-1

Query Match 100.0%; Score 2190; DB 1; Length 404;
Best Local Similarity 100.0%; Pred. No. 1e-228;
Matches 404; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTDDMLHWKMKHNGEKDYSPFSDAEMTRNDVRGMANKNDVALFTSYHCINTYSGWL 60
DB 1 MTDDMLHWKMKHNGEKDYSPFSDAEMTRNDVRGMANKNDVALFTSYHCINTYSGWL 60
QY YCYFGKRYGMVLDHNNATTISAGIDGGQPMRRSFGDNITYTDRRDNFYRAVQLTTGAK 120
DB 61 YCYFGKRYGMVLDHNNATTISAGIDGGQPMRRSFGDNITYTDRRDNFYRAVQLTTGAK 120
QY 121 RIGIEFDHVNLDPRQLERLPGVEFVDISQSPMMRTIKSLBEQKLIREGARVCVGA 180
DB 121 RIGIEFDHVNLDPRQLERLPGVEFVDISQSPMMRTIKSLBEQKLIREGARVCVGA 180
QY 181 ACAAAIKAGVPEHEVAIATTNAMIRESKSPFVELMDTWTFQSGINTDGAHPVTNR1 240
DB 181 ACAAAIKAGVPEHEVAIATTNAMIRESKSPFVELMDTWTFQSGINTDGAHPVTNR1 240
QY 241 VOSGDLISINTFPMIFGYTTALERTLFCDHVDASLDIWEKNVAVHRRGLELIKPGARCK 300
DB 241 VOSGDLISINTFPMIFGYTTALERTLFCDHVDASLDIWEKNVAVHRRGLELIKPGARCK 300
QY 301 DIAIELNEMRWMDLKYRSFGYHSGFYGLCHYYGREAGVELREDIDTELKPGMVSMER 360
DB 301 DIAIELNEMRWMDLKYRSFGYHSGFYGLCHYYGREAGVELREDIDTELKPGMVSMER 360
QY 361 MWMLPEGMFGAGYREHDLIIVGEDAENITGPPFPEHNIIRN 404
DB 361 MWMLPEGMFGAGYREHDLIIVGEDAENITGPPFPEHNIIRN 404

RESULT 2

US-08-947-726A-2
Sequence 2, Application US/08947726A
Patent No. 5932466

GENERAL INFORMATION:

APPLICANT: Furukawa, Keisuke
APPLICANT: Ichikawa, Toshio
APPLICANT: Suzuki, Masaru
APPLICANT: Koyama, Yasuji
TITLE OF INVENTION: A NOVEL CREATIVE AMIDINOHYDROLASE GENE,
TITLE OF INVENTION: A NOVEL RECOMBINANT DNA, AND A PROCESS FOR PRODUCING
TITLE OF INVENTION: CREATIVE AMIDINOHYDROLASE
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDWARDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/947,726A
FILING DATE: 09-OCT-1997
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/535,444

FILING DATE: 27-SEP-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 94/235737
FILING DATE: 29-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lawrence III, Stanton T.
REGISTRATION NUMBER: 25,736
REFERENCE/DOCKET NUMBER: 7005-120-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 404 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-947-726A-2

Query Match 100.0%; Score 2190; DB 2; Length 404;
Best Local Similarity 100.0%; Pred. No. 1e-228;
Matches 404; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTDDMLHWKMKHNGEKDYSPFSDAEMTRNDVRGMANKNDVALFTSYHCINTYSGWL 60
DB 1 MTDDMLHWKMKHNGEKDYSPFSDAEMTRNDVRGMANKNDVALFTSYHCINTYSGWL 60
QY 61 YCYFGKRYGMVLDHNNATTISAGIDGGQPMRRSFGDNITYTDRRDNFYRAVQLTTGAK 120
DB 61 YCYFGKRYGMVLDHNNATTISAGIDGGQPMRRSFGDNITYTDRRDNFYRAVQLTTGAK 120
QY 121 RIGIEFDHVNLDPRQLERLPGVEFVDISQSPMMRTIKSLBEQKLIREGARVCVGA 180
DB 121 RIGIEFDHVNLDPRQLERLPGVEFVDISQSPMMRTIKSLBEQKLIREGARVCVGA 180
QY 181 ACAAAIKAGVPEHEVAIATTNAMIRESKSPFVELMDTWTFQSGINTDGAHPVTNR1 240
DB 181 ACAAAIKAGVPEHEVAIATTNAMIRESKSPFVELMDTWTFQSGINTDGAHPVTNR1 240
QY 241 VOSGDLISINTFPMIFGYTTALERTLFCDHVDASLDIWEKNVAVHRRGLELIKPGARCK 300
DB 241 VOSGDLISINTFPMIFGYTTALERTLFCDHVDASLDIWEKNVAVHRRGLELIKPGARCK 300
QY 301 DIAIELNEMRWMDLKYRSFGYHSGFYGLCHYYGREAGVELREDIDTELKPGMVSMER 360
DB 301 DIAIELNEMRWMDLKYRSFGYHSGFYGLCHYYGREAGVELREDIDTELKPGMVSMER 360
QY 361 MWMLPEGMFGAGYREHDLIIVGEDAENITGPPFPEHNIIRN 404
DB 361 MWMLPEGMFGAGYREHDLIIVGEDAENITGPPFPEHNIIRN 404

RESULT 3

US-08-799-897-1
Sequence 1, Application US/08799897
Patent No. 6080553

GENERAL INFORMATION:

APPLICANT: Sogabe, Atsushi
APPLICANT: Hattori, Takashi
APPLICANT: Nishiyama, Yoshiaki
APPLICANT: Kawamura, Yoshihisa
TITLE OF INVENTION: NOVEL CREATIVE AMIDINOHYDROLASE, PRODUCTION
TITLE OF INVENTION: THEREOF AND USE THEREOF
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: LEYDIG, VOIT & MAYER, LTD.
STREET: Two Prudential Plaza, Suite 4900
CITY: Chicago
STATE: Illinois
COUNTRY: US
ZIP: 60601-6780

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/799,897
; FILING DATE: 13-FEB-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 25435/1996
; FILING DATE: 13-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Robert F. Green
; REGISTRATION NUMBER: 27555
; REFERENCE/DOCKET NUMBER: 78064
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 616-5600
; TELEFAX: (312) 616-5700
; TELEX: 25-3533
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 404 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Alcaligenes faecalis
; STRAIN: TE3581 (FERM P-14237)
; FEATURE:
; NAME/KEY: mat peptide
; LOCATION: 1 to 404
; OTHER INFORMATION: protein having creatine amidino-
; OTHER INFORMATION: hydrolase activity
US-08-799-897-1

Query Match      100.0%; Score 2190; DB 3; Length 404;
Best Local Similarity 100.0%; Pred. No. 1e-228;
Matches 404; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTDDMLHMKMNGEKDYSFSDAEMTRNDVRGMMAKNVDAALFTSYHCINYYSGWL 60
DB 1 MTDDMLHMKMNGEKDYSFSDAEMTRNDVRGMMAKNVDAALFTSYHCINYYSGWL 60
QY 61 YCYFGRKYGWVIDHNNATTISAGIDGQPWRRSFGDNITYTDMRRDNFYRAVQLTTGAK 120
DB 61 YCYFGRKYGWVIDHNNATTISAGIDGQPWRRSFGDNITYTDMRRDNFYRAVQLTTGAK 120
QY 61 YCYFGRKYGWVIDHNNATTISAGIDGQPWRRSFGDNITYTDMRRDNFYRAVQLTTGAK 120
DB 61 YCYFGRKYGWVIDHNNATTISAGIDGQPWRRSFGDNITYTDMRRDNFYRAVQLTTGAK 120
QY 121 RIGIEFDHVNLDPRQLLEALPGVEFVDISQPSMMWRITKLSLEOKLIEGARVCVCGA 180
DB 121 RIGIEFDHVNLDPRQLLEALPGVEFVDISQPSMMWRITKLSLEOKLIEGARVCVCGA 180
QY 121 RIGIEFDHVNLDPRQLLEALPGVEFVDISQPSMMWRITKLSLEOKLIEGARVCVCGA 180
DB 121 RIGIEFDHVNLDPRQLLEALPGVEFVDISQPSMMWRITKLSLEOKLIEGARVCVCGA 180
QY 181 ACAAAIKAGVPEHEVAIATTNMIKREIASSPFEVLMDTWTFQSGINTDGANPVTNRI 240
DB 181 ACAAAIKAGVPEHEVAIATTNMIKREIASSPFEVLMDTWTFQSGINTDGANPVTNRI 240
QY 181 ACAAAIKAGVPEHEVAIATTNMIKREIASSPFEVLMDTWTFQSGINTDGANPVTNRI 240
DB 181 ACAAAIKAGVPEHEVAIATTNMIKREIASSPFEVLMDTWTFQSGINTDGANPVTNRI 240
QY 241 VOSGDILSLNTPFMIFGYTTALERTLFCDHVDASDIWEKNAVARRGLELIKPGARCK 300
DB 241 VOSGDILSLNTPFMIFGYTTALERTLFCDHVDASDIWEKNAVARRGLELIKPGARCK 300
QY 241 VOSGDILSLNTPFMIFGYTTALERTLFCDHVDASDIWEKNAVARRGLELIKPGARCK 300
DB 241 VOSGDILSLNTPFMIFGYTTALERTLFCDHVDASDIWEKNAVARRGLELIKPGARCK 300
QY 301 DIAIELNEMRYREMDLLKTRSPGYGHSFGVLCHYGRAGVLELREDIDTELKPGMVVSMEP 360
DB 301 DIAIELNEMRYREMDLLKTRSPGYGHSFGVLCHYGRAGVLELREDIDTELKPGMVVSMEP 360
QY 301 DIAIELNEMRYREMDLLKTRSPGYGHSFGVLCHYGRAGVLELREDIDTELKPGMVVSMEP 360
DB 301 DIAIELNEMRYREMDLLKTRSPGYGHSFGVLCHYGRAGVLELREDIDTELKPGMVVSMEP 360
QY 361 MWMLPEGMGAGGYRHHDLIIYGEDGAENITGPPGPEHHIIRN 404
DB 361 MWMLPEGMGAGGYRHHDLIIYGEDGAENITGPPGPEHHIIRN 404

RESULT 4
US-09-856-645-1
; Sequence 1, Application US/09856645
```

```

; Patent No. 6639700
; GENERAL INFORMATION:
; APPLICANT: KIKKOMAN CORPORATION
; TITLE OF INVENTION: CREATINE AMIDINOHYDROLASE AND PROCESSES FOR PRODUCING THE SAME
; FILE REFERENCE: PH-688-PCT
; CURRENT APPLICATION NUMBER: US/09/856,645
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: JP10-334252
; PRIOR FILING DATE: 1998-11-25
; NUMBER OF SEQ ID NOS: 1
; SEQ ID NO 1
; LENGTH: 404
; TYPE: PRT
; ORGANISM: Alcaligenes sp.
US-09-856-645-1

Query Match      100.0%; Score 2190; DB 4; Length 404;
Best Local Similarity 100.0%; Pred. No. 1e-228;
Matches 404; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTDDMLHMKMNGEKDYSFSDAEMTRNDVRGMMAKNVDAALFTSYHCINYYSGWL 60
DB 1 MTDDMLHMKMNGEKDYSFSDAEMTRNDVRGMMAKNVDAALFTSYHCINYYSGWL 60
QY 61 YCYFGRKYGWVIDHNNATTISAGIDGQPWRRSFGDNITYTDMRRDNFYRAVQLTTGAK 120
DB 61 YCYFGRKYGWVIDHNNATTISAGIDGQPWRRSFGDNITYTDMRRDNFYRAVQLTTGAK 120
QY 61 YCYFGRKYGWVIDHNNATTISAGIDGQPWRRSFGDNITYTDMRRDNFYRAVQLTTGAK 120
DB 61 YCYFGRKYGWVIDHNNATTISAGIDGQPWRRSFGDNITYTDMRRDNFYRAVQLTTGAK 120
QY 121 RIGIEFDHVNLDPRQLLEALPGVEFVDISQPSMMWRITKLSLEOKLIEGARVCVCGA 180
DB 121 RIGIEFDHVNLDPRQLLEALPGVEFVDISQPSMMWRITKLSLEOKLIEGARVCVCGA 180
QY 121 RIGIEFDHVNLDPRQLLEALPGVEFVDISQPSMMWRITKLSLEOKLIEGARVCVCGA 180
DB 121 RIGIEFDHVNLDPRQLLEALPGVEFVDISQPSMMWRITKLSLEOKLIEGARVCVCGA 180
QY 181 ACAAAIKAGVPEHEVAIATTNMIKREIASSPFEVLMDTWTFQSGINTDGANPVTNRI 240
DB 181 ACAAAIKAGVPEHEVAIATTNMIKREIASSPFEVLMDTWTFQSGINTDGANPVTNRI 240
QY 181 ACAAAIKAGVPEHEVAIATTNMIKREIASSPFEVLMDTWTFQSGINTDGANPVTNRI 240
DB 181 ACAAAIKAGVPEHEVAIATTNMIKREIASSPFEVLMDTWTFQSGINTDGANPVTNRI 240
QY 241 VOSGDILSLNTPFMIFGYTTALERTLFCDHVDASDIWEKNAVARRGLELIKPGARCK 300
DB 241 VOSGDILSLNTPFMIFGYTTALERTLFCDHVDASDIWEKNAVARRGLELIKPGARCK 300
QY 241 VOSGDILSLNTPFMIFGYTTALERTLFCDHVDASDIWEKNAVARRGLELIKPGARCK 300
DB 241 VOSGDILSLNTPFMIFGYTTALERTLFCDHVDASDIWEKNAVARRGLELIKPGARCK 300
QY 301 DIAIELNEMRYREMDLLKTRSPGYGHSFGVLCHYGRAGVLELREDIDTELKPGMVVSMEP 360
DB 301 DIAIELNEMRYREMDLLKTRSPGYGHSFGVLCHYGRAGVLELREDIDTELKPGMVVSMEP 360
QY 301 DIAIELNEMRYREMDLLKTRSPGYGHSFGVLCHYGRAGVLELREDIDTELKPGMVVSMEP 360
DB 301 DIAIELNEMRYREMDLLKTRSPGYGHSFGVLCHYGRAGVLELREDIDTELKPGMVVSMEP 360
QY 361 MWMLPEGMGAGGYRHHDLIIYGEDGAENITGPPGPEHHIIRN 404
DB 361 MWMLPEGMGAGGYRHHDLIIYGEDGAENITGPPGPEHHIIRN 404

RESULT 5
US-09-869-280A-1
; Sequence 1, Application US/09869280A
; GENERAL INFORMATION:
; APPLICANT: Koykawa, Keisuke
; APPLICANT: Koykawa, Yasuji
; APPLICANT: Suzuki, Masaru
; TITLE OF INVENTION: THERMOSTABLE CREATINE AMIDINOHYDROLASE AND PROCESS FOR
; FILE REFERENCE: 04853-0075-00000
; CURRENT APPLICATION NUMBER: US/09/869,280A
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: PCT/JP99/07424
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: JP-11-033359
; PRIOR FILING DATE: 1999-01-01
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 1
; LENGTH: 404
; TYPE: PRT
; ORGANISM: Alcaligenes sp.
US-09-869-280A-1
```


CURRENT FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/107,433
PRIOR FILING DATE: 1998-06-30
PRIOR APPLICATION NUMBER: US 60/085,131
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: US 60/051,553
PRIOR FILING DATE: 1997-07-02
NUMBER OF SEQ ID NOS: 5322
SEQ ID NO 3285
LENGTH: 360
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-583-110-3285

Query Match 8.6%; Score 188.5; DB 4; Length 360;
Best Local Similarity 23.9%; Fred. No. 1.3e-11;
Matches 76; Conservative 50; Mismatches 139; Indels 53; Gaps 11;

QY 84 IDGGPWRRSFGDNTTYTDMRDNFRAVROLTTGAKRIGIEFDHNLDFRQLBEALPG 143
DB 76 VDSNPWOK-----IKHALPOL--DFKRAVEFDLILTKHGLTKVFEET 118
QY 144 VEFVDISOPMMWRITKSLBE-OKLIREGARVCDVGAAACAIAKGV-----PEHEV 195
DB 119 AEFNDLTPRIQRRLIKSADEVQKM-----VAGLYADKXAVHGFNDISLDTKETDI 170
QY 196 AIAATNAMIRE-IAKSPFVELMDTWTWFGSINTDGAH-PYTKRIYOSGDIISLNTFP 253
DB 171 IAQIDPAMKRREGVMSFD-----TWVLTDGNANPHGIPPAANK-VENDALLLFDLGV 221
QY 254 MIFGYTALERTLPCDHVDASLDIWEKNVAVHRRGLEIKPGARCKOIAELNEMVREM 313
DB 222 LVNGYASDMTRIVAVGKPDQFKDILNLTLEAQNALDFIKGVTAAHEVDRAAREVIEKA 281
QY 314 DLKTRSPGYGSHFGLCHYGRAGVELREDIDTELKPGVMSVPMVMLPEGMPGAG 373
DB 282 GGYEYFNRLHGHIGMDVHEF-----PSIMEGNDVIEGMCFSVEPGIYI----PGKVG 332
QY 374 YREHDILIVGEDGAENIT 391
DB 333 VRIEDCGVVTXKGFDLFT 350

RESULT 11

US-09-107-433-3390
Sequence 3390, Application US/09107433
Patent No. 6800744
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
THERAPEUTICS
NUMBER OF SEQUENCES: 5206
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: <Unknown>
OPERATING SYSTEM: <Unknown>
SOFTWARE: <Unknown>
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,433
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/ 085131
FILING DATE: May 12, 1998
APPLICATION NUMBER: 60/051553
FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-011
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 3390:
SEQUENCE CHARACTERISTICS:
LENGTH: 361 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Streptococcus pneumoniae
FEATURE:
NAME/KEY: misc_feature
LOCATION: (B) LOCATION 1..361
SEQUENCE DESCRIPTION: SEQ ID NO: 3390:
US-09-107-433-3390

Query Match 8.6%; Score 188.5; DB 4; Length 361;
Best Local Similarity 23.9%; Fred. No. 1.3e-11;
Matches 76; Conservative 50; Mismatches 139; Indels 53; Gaps 11;

QY 84 IDGGPWRRSFGDNTTYTDMRDNFRAVROLTTGAKRIGIEFDHNLDFRQLBEALPG 143
DB 77 VDSNPWOK-----IKHALPOL--DFKRAVEFDLILTKHGLTKVFEET 119
QY 144 VEFVDISOPMMWRITKSLBE-OKLIREGARVCDVGAAACAIAKGV-----PEHEV 195
DB 120 AEFNDLTPRIQRRLIKSADEVQKM-----VAGLYADKXAVHGFNDISLDTKETDI 171
QY 196 AIAATNAMIRE-IAKSPFVELMDTWTWFGSINTDGAH-PYTKRIYOSGDIISLNTFP 253
DB 172 IAQIDPAMKRREGVMSFD-----TWVLTDGNANPHGIPPAANK-VENDALLLFDLGV 222
QY 254 MIFGYTALERTLPCDHVDASLDIWEKNVAVHRRGLEIKPGARCKOIAELNEMVREM 313
DB 223 LVNGYASDMTRIVAVGKPDQFKDILNLTLEAQNALDFIKGVTAAHEVDRAAREVIEKA 282
QY 314 DLKTRSPGYGSHFGLCHYGRAGVELREDIDTELKPGVMSVPMVMLPEGMPGAG 373
DB 283 GGYEYFNRLHGHIGMDVHEF-----PSIMEGNDVIEGMCFSVEPGIYI----PGKVG 333
QY 374 YREHDILIVGEDGAENIT 391
DB 334 VRIEDCGVVTXKGFDLFT 351

RESULT 12

US-09-602-777A-94
Sequence 94, Application US/09602777A
Patent No. 6831165
GENERAL INFORMATION:
APPLICANT: Pompeius, Markus
APPLICANT: Kroger, Burkhard
APPLICANT: Schroder, Hartwig
APPLICANT: Zelder, Oskar
APPLICANT: Haberer, Gregor
TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
INVOLVED IN HOMEOSTASIS AND ADAPTATION
FILE REFERENCE: BGI-128CP
CURRENT APPLICATION NUMBER: US/09/602,777A
PRIOR FILING DATE: 2000-06-23
PRIOR APPLICATION NUMBER: US 60/141031
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: DE 19931636.8
PRIOR FILING DATE: 1999-07-08
PRIOR APPLICATION NUMBER: DE 19932125.6
PRIOR FILING DATE: 1999-07-09
PRIOR APPLICATION NUMBER: DE 19932126.4

```

; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932127.2
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932128.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932129.9
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: DE 19932226.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932920.6
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932922.2
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932924.9
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932930.3
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932932.8
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932935.4
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932973.7
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933002.6
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933003.4
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933005.0
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933006.9
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19941378.9
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941379.7
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941390.8
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941391.6
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19942088.2
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 442
; SEQ ID NO 94
; LENGTH: 363
; TYPE: PRP
; ORGANISM: Corynebacterium glutamicum
US-09-602-777A-94

```

```

Query Match      8.6%; Score 187.5; DB 4; Length 363;
Best Local Similarity 22.2%; Pred. No. 1.7e-11;
Matches 91; Conservative 56; Mismatches 164; Indels 99; Gaps 12;

QY 22 SPAEMTRNDVRCGMAKKNVDAALFTSHCINYYSGWL-----YCYFGR 66
DB 4 ADTRPATRRALAAKLAQRIIDILVTSPIHAYILSGFGSCALIVKDLASQITDGR 63
QY 67 KYGMV-----IDHNNATTISAGIDGQFWRRSFGDNITTYTDRRDNFYRAVROLT 116
DB 64 YTTQIAEVPDIEALIERASATTLAQVE----- 92
QY 117 TCAKRIIGIFDHVNLDFRROLEAL-PGVEFVDISOPSMWRTIKSLEROKLIREGARVC 175
DB 93 -CPRRIRAIBAQTTLQDLISREATQEDVELIPVS-----GVVESIRLTKOSFELDRLR 145
QY 176 DVGGA-----CAAIRKAGVBEHVAIATNAMIIRIAXSPFVELMDTWTFQSGI 227
DB 146 DVAALASQAFEDLLAAGEIABEGSEROVA-ADLETYRRLIGARBPFD-----TIYASGP 199
QY 228 NTDGANPNYTNRIVOSGDILSLNTFPMIGYATALERTLFCDHVDASLDIWEKNVAHR 287
DB 200 NAKXPHHGAGDRIILRGDGLVTIDFGAHAGFNSDMRTITLWMGEAGFEBAIYDIYLRSQL 259

```

```

QY 288 RLELEIKPGARCKDIAIEINEMYREMDLKYRSFGYHSGFVLCHYYGREGAVELREDID 347
DB 260 AGVEAAYGANLFDIDAACKRIED-----AGYGEF---VHSTHGIGIGLEVHEARS 308
QY 348 TE-----LKRGVVSMEBPMVLPFGCMGAGCYREHDIIVGEDGAENIT 391
DB 309 ASKTSQGLVLETSSTLIEPGIY---VPGKGVRIEDTLIYSGAPEIIT 354

RESULT 13
US-09-583-110-4793
; Sequence 4793, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 4793
; LENGTH: 353
; TYPE: PRP
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-4793

```

```

Query Match      8.4%; Score 183; DB 4; Length 353;
Best Local Similarity 24.4%; Pred. No. 5.1e-11;
Matches 71; Conservative 55; Mismatches 121; Indels 44; Gaps 11;

QY 118 GAKRIGIFEDHVNLDPRROLEALPGVEFVDISOPSMWRTIKSLEROKLIREGARVCV 177
DB 87 GLTRIGFE-DEISVYTHRMQAAHFGGLDFPOTQVVEGIRMTKDEEYIAIKACISIQ 145
QY 178 GGAACAAMIKAGVBEHVAIATNAMIIRIAXSPFVELMDTWTFQSGINTGAPVY 237
DB 146 AFRDALDPFKPKETIEIA-NFLDPRMELASG-----LSPTDILASGINSKPHAHM 199
QY 228 NRIVOSGDILSLNTFPMIFGYTTA-LERTLFCDHVDASLDIWEKNVAHRGELIKFG 296
DB 200 HKPVELGERITWD-FGCIYDHYVSDMTRTIYLGHVSDQAEIYNTVLKANQALIDAKAG 258
QY 297 ARC-----KDIAIENEMYREMDLKYRSFGYHSGFVLCHYYGREGAVELREDID 345
DB 259 IGFRRDFDKIPRIITIEA-----GYGDYF---THGIGHGIGLDIHEPPIYS 300
QY 346 -IDTE-LKRGVVSMEBPMVLPFGCMGAGCYREHDIIVGEDGAENITGP 394
DB 301 QSTETIKGVALTDEPGIYI---EGKKGVAIEDDILITRGCELLTAP 347

RESULT 14
US-09-107-433-3157
; Sequence 3157, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham

```

```

STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: <Unknown>
OPERATING SYSTEM: <Unknown>
SOFTWARE: <Unknown>
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,433
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085131
FILING DATE: May 12, 1998
APPLICATION NUMBER: 60/051553
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Arinello, Pamela Dencke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-011
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 3157:
SEQUENCE CHARACTERISTICS:
LENGTH: 358 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Streptococcus pneumoniae
FEATURE:
NAME/KEY: misc feature
LOCATION: (8) LOCATION 1...358
SEQUENCE DESCRIPTION: SEQ ID NO: 3157:
US-09-107-433-3157

Query Match      8.4%; Score 183; DB 4; Length 358;
Best Local Similarity 24.4%; Pred. No. 5.2e-11;
Matches 71; Conservative 55; Mismatches 121; Indels 44; Gaps 11;

QY 118 GAKRIGIEPDHVNLDPRQLEALPGVEFVDISQPSMMRTIKSLSEQKLIREGARVCDV 177
   ||| ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 92 GLTRIGFE-DEISVSYYHMOAFAGLDLPQTFVEGLRMIXDEALAIKACSIDQ 150
   ||| ||| : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 178 GGAACAATKAGVEHEVAIATTNAMIKEIAKSPFVELMDTWTWFOGINTDGAHPVT 237
   ||| ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 151 AFRDALDFIKPKETLEIA-NFLDFRRELGAAG-----LSFDTLIASSGINSKPHAPM 204
   ||| ||| : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 238 NRIVSGDILLNTPPMIFGYTA-LERTLPCHVDASLDIWEKNVAVHRGGLLEIKPG 296
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 205 HKPVELGEALITMD-FCGLVDHYSDMTRTIYGHVSDQEALEYNTVLKANQALIDQAKAG 263
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 297 ARC-----KQIAIFLNEMREMDLKTYSFGYGHSPGLCHYYREAGVELARE----- 345
   ||| ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 264 LGFRDFDKLPRIIIEA-----GYGDYF---TIGIGHGIGLDIHEEPYFS 305
   ||| ||| : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 346 -IDTE-LKPGVVSMEPPVMLPEGMPEGAGYREHDLIVGEDGAENITGFP 394
   ||| ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 306 QSTETIKTKGMALTDPEGIY---EGKGVRIEDDILITETGCCHLTAP 352
   ||| ||| : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 15
US-09-861-451A-36
Sequence 36, Application US/09861451A
Patent No. 6759316
GENERAL INFORMATION:
APPLICANT: Commonwealth Scientific & Industrial Research Orga
TITLE OF INVENTION: Methods of Identifying Antigen Gene Sequences
FILE REFERENCE: PF34033/01
CURRENT APPLICATION NUMBER: US/09/861,451A
CURRENT FILING DATE: 2001-05-21

```

```

PRIOR APPLICATION NUMBER: PF7273
PRIOR FILING DATE: 1998-11-20
NUMBER OF SEQ ID NOS: 84
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 36
LENGTH: 299
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Deduced protein
US-09-861-451A-36

Query Match      8.3%; Score 181.5; DB 4; Length 299;
Best Local Similarity 22.5%; Pred. No. 5.7e-11;
Matches 61; Conservative 49; Mismatches 132; Indels 29; Gaps 6;

QY 127 DHVNDLPRQLEALPGVEFVDISQPSMMRTIKSLSEQKLIREGARVCDVGAACAAT 186
   : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 46 EYLITDQDKIRSWFPNADPVKIQ--AQLPRLIKTEEIKNIKXAVEISLAIVNKIPKI 103
   : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 187 KAGVEHEVAIATTNAMIKEIAKSPFVELMDTWTWFOGINTDGAHPVTNRIVQSGDI 246
   ||| ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 104 KPGMTEKSIDVNLNQMKLGAEKESFDSIIAT-----GNSAMPHWRASETEILDNDL 157
   ||| ||| : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 247 LSLNTPPMIFGYTALERTLPCHVDASLDIWEKNVAVHRGGLLEIKPGARCKQIATL 306
   : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 158 LKIDGALFNQYCADITFTSYIGQISEKDLILEIKVKAELIGKKVAPGVKASEIDLAC 217
   : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 307 NEMVREMDLKYRSPGYSHSPGLCHYYREAGVELARE-----DIDTELKPGVVSMEP 360
   ||| ||| : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 218 RNFTIEQ-----GYKVF---IHSTGHGVGIDIHLLPVVSSSTQTLERGVITVER 266
   ||| ||| : : : : : : : : : : : : : : : : : : : : : : : : : :

QY 361 VVMLPEGMPEGAGYREHDLIVGEDGAENIT 391
   : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 267 GIYI---PGLGARIEDVVLVTESGFRTLS 293
   : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Search completed: July 8, 2005, 01:36:46
 Job time : 29 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 8, 2005, 01:35:47 ; Search time 77 Seconds

(without alignments)
2026.980 Million cell updates/sec

Title: US-10-807-228a-1
Perfect score: 2190
Sequence: 1 MTDDMLHMKHNGEKDYSP.....DGAENITGPPRPHENITRN 404

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1726216 seqs, 386330316 residues

Total number of hits satisfying chosen parameters: 1726216

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
19: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep:*
20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2078	94.9	404	US-10-251-078-2	Sequence 2, Appli
2	2078	94.9	404	US-10-251-078-18	Sequence 10, Appl
3	2068	94.4	404	US-10-251-078-10	Sequence 10, Appl
4	2068	94.4	404	US-10-251-078-22	Sequence 22, Appl
5	2066	94.3	404	US-10-251-078-16	Sequence 16, Appl
6	2064	94.2	404	US-10-251-078-12	Sequence 12, Appl
7	2062	94.2	404	US-10-251-078-14	Sequence 14, Appl
8	2058	94.0	404	US-10-251-078-20	Sequence 20, Appl
9	248.5	10.5	354	US-10-369-493-21632	Sequence 21632, A
10	229.5	10.5	354	US-09-815-242-10645	Sequence 10645, A
11	229	10.5	361	US-09-815-242-10235	Sequence 10235, A

	12	203.5	9.3	381	14	US-10-156-761-9699	Sequence 9699, Ap
	13	200	9.1	350	9	US-09-815-242-5287	Sequence 5287, Ap
	14	200	9.1	353	9	US-09-815-242-12580	Sequence 12580, A
	15	194.5	8.9	353	15	US-10-282-122A-70734	Sequence 70734, A
	16	190.5	8.7	360	17	US-10-472-928-3260	Sequence 3260, Ap
	17	189	8.6	353	9	US-09-815-242-13193	Sequence 13193, A
	18	188.5	8.6	361	18	US-10-617-320-3390	Sequence 3390, Ap
	19	187.5	8.6	363	9	US-09-738-626-5290	Sequence 5290, Ap
	20	187.5	8.6	368	14	US-10-156-761-14386	Sequence 14386, A
	21	183	8.4	353	17	US-10-472-928-184	Sequence 184, Ap
	22	183	8.4	358	18	US-10-617-320-3157	Sequence 3157, Ap
	23	181.5	8.3	299	9	US-09-861-451A-36	Sequence 36, Appl
	24	179	8.2	243	10	US-09-877-843-94	Sequence 94, Appl
	25	171	7.8	355	15	US-10-289-762-871	Sequence 871, Appl
	26	166	7.6	240	15	US-10-042-865-189	Sequence 189, Appl
	27	162.5	7.4	375	16	US-10-482-706-141	Sequence 141, Appl
	28	160.5	7.3	253	15	US-10-289-867-10	Sequence 10, Appl
	29	160.5	7.3	253	15	US-10-350-516-4	Sequence 4, Appl
	30	160.5	7.3	397	15	US-10-424-599-247413	Sequence 247413, A
	31	160	7.3	352	15	US-10-282-122A-71551	Sequence 71551, A
	32	158.5	7.2	350	16	US-10-425-115-369034	Sequence 369034, A
	33	156.5	7.1	306	15	US-10-425-114-58168	Sequence 58168, A
	34	155	7.1	428	15	US-10-369-493-245	Sequence 245, Appl
	35	152.5	7.0	378	15	US-10-369-493-16426	Sequence 16426, A
	36	151.5	6.9	417	15	US-10-369-493-10239	Sequence 10239, A
	37	151	6.9	391	16	US-10-425-115-299292	Sequence 299292, A
	38	146.5	6.7	437	15	US-10-369-493-21222	Sequence 21222, A
	39	145	6.6	356	15	US-10-424-599-173546	Sequence 173546, A
	40	143.5	6.6	369	9	US-10-042-865-186	Sequence 186, Appl
	41	143	6.5	379	9	US-09-738-626-5142	Sequence 5142, Appl
	42	142.5	6.5	369	15	US-10-042-865-187	Sequence 187, Appl
	43	142	6.5	274	15	US-10-299-867-11	Sequence 11, Appl
	44	141	6.4	432	16	US-10-739-930-9633	Sequence 9633, Appl
	45	140.5	6.4	248	15	US-10-299-867-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1
US-10-251-078-2
; Sequence 2, Application US/10251078
; Publication No. US200301119084A1
; GENERAL INFORMATION:
; APPLICANT: Roche Diagnostics GmbH
; APPLICANT: Shao, Zhixin
; APPLICANT: Schmuck, Rainer
; APPLICANT: Kratzsch, Peter
; APPLICANT: Kenkies, Janet
; APPLICANT: Weisner, Harald
; TITLE OF INVENTION: Variants of an Erwinia-type creatinase
; FILE REFERENCE: 20981 EP
; CURRENT APPLICATION NUMBER: US/10/251.078
; CURRENT FILING DATE: 2002-09-20
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 2
; LENGTH: 404
; TYPE: PRT
; ORGANISM: Erwinia sp.
US-10-251-078-2

Query Match	Score	2078	DB 14	Length	404
Best Local Similarity	94.3%	Pred. No.	4.1e-195		
Matches	381	Conservative	14	Mismatches	9
				Indels	0
				Gaps	0

QY 1 MTDDMLHMKHNGEKDYSPSDAEMTRQNDVGRMAKNNVDAALFTSYCINYSGL 60
DB 1 MTDDMLHMKHNGEKDYSPSDAEMTRQNDVGRMAKNNVDAALFTSYCINYSGL 60
QY 61 YCYGRKKYGVITDNNATTTISAGIDGQPMRRSRFDNITTYTDWRDNNFRAVROLTTGAK 120
DB 61 YCYGRKKYGVITDNNATTTISAGIDGQPMRRSRFDNITTYTDWRDNNFRAVROLTTGAK 120

Qy	181	ACAAAIKAGVPHEHEVALATNTAMIRIAKSPPEVELMDTWTFQSGINDGANNPTNRI	240
Db	121	RIGIEFDHVNIDFRRLOLEALPGVEFVDSIQSSMMRRITKSJLEEQKLIIEGARVCVGA	180
Qy	121	RIGIEFDHVNIDFRRLOLEALPGVEFVDSIQSSMMRRITKSJLEEQKLIIEGARVCVGA	180
Db	121	RIGIEFDHVNIDFRRLOLEALPGVEFVDSIQSSMMRRITKSJLEEQKLIIEGARICVGA	180
Qy	181	ACVAALKAGVPHEHEVALATNTAMIRIAKSPPEVELMDTWTFQSGINDGANNPTNRI	240
Db	181	ACVAALKAGVPHEHEVALATNTAMIRIAKSPPEVELMDTWTFQSGINDGANNPTNRI	240
Qy	241	VQSGIISLNTPEMIFGYTALERTFCCHVDASLDIMEKNAVRRRGELIKPGARCK	300
Db	241	VQSGIISLNTPEMIFGYTALERTFCCHVDASLDIMEKNAVRRRGELIKPGARCK	300
Qy	301	DIAIEINEMYRMDILKTRSPFGYGSFGYLCHYYGEAGVELREDIDTELKPGMVSMEP	360
Db	301	DIAIEINEMYRMDILKTRSPFGYGSFGYLCHYYGEAGVELREDIDTELKPGMVSMEP	360
Qy	361	MYMLPEGMGAGGYREHDLIIYGEDEAKENITGPPPENNIIRN	404
Db	361	MYMLPEGMGAGGYREHDLIIYGEDEAKENITGPPPENNIIRN	404

```

RESULT 2
US-10-251-078-18
: Sequence 18, Application US/10251078
: Publication NO. US20030119084A1
: GENERAL INFORMATION:
: APPLICANT: Roche Diagnostics GmbH
: APPLICANT: Shao, Zhixin
: APPLICANT: Schmuck, Rainer
: APPLICANT: Kratzsch, Peter
: APPLICANT: Kenkies, Janet
: APPLICANT: Welser, Harald
: TITLE OF INVENTION: Variants of an Erythrina-type creatinase
: FILE REFERENCE: 20981 BP
: CURRENT APPLICATION NUMBER: US/10/251, 078
: CURRENT FILING DATE: 2002-09-20
: NUMBER OF SEQ ID NOS: 32
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 18
: LENGTH: 404
: TYPE: PR1
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence:variant C1gc2
US-10-251-078-18

          94.9%; Score 2078; DB 14; Length 404;
Query Match      Beel Local Similarity 94.1%; Pred. No. 4,1e-195;
Matches 380; Conservative 14; Mismatches 10; Indels 0; Gaps 0

QY      1      MTDMLAHMKWNGEKEDYSPFSDAEMTRONDVRCGMANKNVDAALFTSYHCINYSGL 60
DB      1      MTDMLAHMKWNGEKEXSPFSDAEMTRQSDVRRMAENDVDAALFTSYHCINYSGL 60
QY      61      YCYFGRKKGWYIDHNNATTISAGIDGGOPRRSPFGDNTTYDWRKDNFTYRAVRLTGGK 120
DB      61      YCYFGRKKGWYIDOHATTISAGIDGGOPRRSPFGDNTTYDWRKDNFTYRAVRLTGGK 120
QY      121      RIGIEFDHVNIDFRQLBEALPGVFEVDISOPSMRRTKSLSEQGLREGARVCDVGGA 180
DB      121      RIGIEFDHVDLDFKRTLEBALGVFEVDIGQSMRRTVKSLEQGLREGARKCDVGGA 180
QY      181      ACAAIKKGVEHEVAIAITTNMIREIAKSPFEVLEMDTWTFQSGINTDGAHNVTRNI 240
DB      181      ACVAAYKGVGEHEVAIAITTNMAREIAKSPFEVLEMDTWTFQSGINTDGAHNVTRNI 240
QY      241      VQSGDIISLNTFPMIFGYTALERTLPCDHVDASLDIWEKVAVHRRGLEIKGGRCK 300
DB      241      VQSGDIISLNTFPMIFGYTALERTLPCDHVDASLDIWEKVAVHRRGLEIKGGRCK 300
QY      301      DIAELNMEVEMDLKTKRSPCYGSGFGLCHYYGREGVLEKEDIDLELKCGMWVSMER 360
DB      301      DIAELNMEVEMDLKTKRSPCYGSGFGLCHYYGREGVLEKEDIDLELKCGMWVSMER 360

```

Db 301 DIALIENMYRRMDDLLKTRSFCEYCHSFVLSHYHGRAGVGLRBDIDTVLQPGNVISMSP 360

Qy 361 MWMLPEGMPAGGYREHDILLVGEDEAGNIITGPPGPEHNIIRN 404

Db 361 MWMLPEGAGPGAGYREHDILLVGEDEAGNIITGPPGPEHNIIRN 404

```

RESULT 3
US-10-251-078-10
Sequence 10. Application US/10251078
Publication No. US20030119084A1
GENERAL INFORMATION:
APPLICANT: Roche Diagnostics GmbH
APPLICANT: Shao, Zhixin
APPLICANT: Schumack, Rainer
APPLICANT: Kratzsch, Peter
APPLICANT: Kenkies, Janet
APPLICANT: Weisser, Harald
TITLE OF INVENTION: Variants of an Erwinia-type creatinase
FILE REFERENCE: 20961 EP
CURRENT APPLICATION NUMBER: US/10/251,078
CURRENT FILING DATE: 2002-09-20
NUMBER OF SEQ. ID NOS: 32
SOFTWARE: PatentIn version 3.1
SEQ ID NO 10
LENGTH: 404
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of variant C1M24
US-10-251-078-10

```

```

Query Match          94.4%; Score 2068; DB 14; Length 404;
Best Local Similarity 93.8%; Pred. No. 3.9e-194;
Matches 379; Conservative 15; Mismatches 10; Indels 0; Gaps 0

QY      1 MTDDMLHVMKHNCKEKOYSPSPDAEMTRKQNDVGMKAKNNVDALFTSYHCINYYSGWL 60
DB      1 MTDDMLHVMKHNCKEKEYSPSPDAEMTRKROSDVRRMAENVDALFTSYHCINYYSGFL 60
QY      61 YCYGRKKXGWIIDHNNATTTISAGIDGGQPMWRSPFGDNITTYTDMRDRNFYRAVROLTTGAK 120
DB      61 YCYGRKKXGWIIDDDHATTTISAGIDGGQPMWRSPFGDNITTYTDMRDRNFYCAVROLTTGAR 120
QY      121 RIGIEFDHVLNDFPRROLEALPGVEFYVDISQPSMMRTIKSLEROKLIREGARVCDVGGA 180
DB      121 RIGIEFDHVLNDFPRRTLEALPGVEFYVDIGQPSMMRTIKSLEROKLIREGARICDVGGA 180
QY      181 ACAAAIAQGVDEHEVALATTNAMIRELAKSPPEFLMDTWTWTFOSGINTDGAHNVTMRI 240
DB      181 ACVAAYAVAGDEHEVALATTNAMIRELAKSPPEFLMDTWTWTFOSGINTDGAHNVTMRI 240
QY      241 VQSGDIIISLTFPMIFGYTTLERTLFCDDHVDASLDIWEKQVAVHRRGLEIKRGARCK 300
DB      241 VQSGDIIISLTFPMIFGYTTLERTLFCDDHVDASLDIWEKQVAVHRRGLEIKRGARCK 300
QY      301 DIAIEINMYREWDLLKTRSPGYSHSPGVLCHYYGREAGVELREDIDTELKPMYVSNBP 360
DB      301 DIAIEINMYREWDLLKTRSPGYSHSPGVLCHYYGREAGVELREDIDTELKPMYVSNBP 360
QY      361 MVMLEPGMFGAGYREHDIILVGEDGAEINTGPPGPEHNIIRN 404
DB      361 MVMLEPGAGAGYREHDIILVGEDGAEINTGPPGPEHNIIRN 404

RESULT 4
US-10-251-078-22
; Sequence 22, Application US/10251078
; Publication No. US20030119084v1
; GENERAL INFORMATION:
; APPLICANT: Roche Diagnostics GmbH
; APPLICANT: Shao, Zhixin
; APPLICANT: Schmuck, Rainer

```

APPLICANT: Kratzsch, Peter
APPLICANT: Kenkies, Janet
TITLE OF INVENTION: Variants of an Erwinia-type creatinase
FILE REFERENCE: 20981 EP
CURRENT APPLICATION NUMBER: US/10/251,078
CURRENT FILING DATE: 2002-09-20
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PatentIn version 3.1
SEQ ID NO 22
LENGTH: 404
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:variant CT2m7
US-10-251-078-22

Query Match 94.4%; Score 2068; DB 14; Length 404;
Best Local Similarity 93.6%; Pred. No. 3.9e-194;
Matches 378; Conservative 15; Mismatches 11; Indels 0; Gaps 0;

QY 1 MTDDMLHMKHNGEKDYSPPSDAEMTRQNDVRGMKANNVDALFTSYHCINYSGL 60
DB 1 MTDDMLHMKHNGEKESPPSDAEMTRQSDVRMMAMENDVDALFTSYHCINYSGL 60
QY YCYFGRKGYMDHNNATTISAGIDGQPMRRSGDNTTYTDMRDNFYRAVROLTPGAK 120
DB 61 YCYFGRKGYMDIDDHATTISAGIDGQPMRRSGDNTTYTDMRDNFYRAVROLTPGAK 120
QY 121 RIGIEFDHVNLDPRQLEALPGVEFVDISQPSMMRTIKSLSEOKLIREGARVCDVGA 180
DB 121 RIGIEFDHVDLDFRRTLEALPGVEFVDISQPSMMRTIKSLSEOKLIREGARVCDVGA 180
QY 181 ACAAIKAGVBEHEVAIATTNAMIKEIKSPFVELMDTWTWFOGINTDGAHNPVTNRI 240
DB 181 ACVAAVKAGVBEHEVAIATTNAMIKEIKSPFVELMDTWTWFOGINTDGAHNPVTNRI 240
QY 241 VOSGDILSLNTPFMIFGYTTALERTLPCDHVDASLDITWEKNVAHRRGLEIKPGARCK 300
DB 241 VOSGDILSLNTPFMIFGYTTALERTLPCDHVDASLDITWEKNVAHRRGLEIKPGARCK 300
QY 301 DIAELNEMRYEMWDLKTRSPGYGHSFGVLCHYYGREAGVELREDIDTELKPGMYVSMEP 360
DB 301 DIAELNEMRYEMWDLKTRSPGYGHSFGVLCHYYGREAGVELREDIDTELKPGMYVSMEP 360
QY 361 MVMLEPGMAGAGYREHDLIVGEDGAENITGPPGPEHNIIRN 404
DB 361 MVMLEPGMAGAGYREHDLIVGEDGAENITGPPGPEHNIIRN 404

RESULT 5
US-10-251-078-16
Sequence 16, Application US/10251078
Publication No. US20030119084A1
GENERAL INFORMATION:
APPLICANT: Roche Diagnostics GmbH
APPLICANT: Shao, Zhixin
APPLICANT: Schmuck, Rainer
APPLICANT: Kratzsch, Peter
APPLICANT: Kenkies, Janet
APPLICANT: Weisner, Harald
TITLE OF INVENTION: Variants of an Erwinia-type creatinase
FILE REFERENCE: 20981 EP
CURRENT APPLICATION NUMBER: US/10/251,078
CURRENT FILING DATE: 2002-09-20
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PatentIn version 3.1
SEQ ID NO 16
LENGTH: 404
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:variant CT2m8

US-10-251-078-16
Query Match 94.3%; Score 2066; DB 14; Length 404;
Best Local Similarity 93.6%; Pred. No. 6.2e-194;
Matches 378; Conservative 16; Mismatches 10; Indels 0; Gaps 0;

QY 1 MTDDMLHMKHNGEKDYSPPSDAEMTRQNDVRGMKANNVDALFTSYHCINYSGL 60
DB 1 MTDDMLHMKHNGEKESPPSDAEMTRQSDVRMMAMENDVDALFTSYHCINYSGL 60
QY 61 YCYFGRKGYMDHNNATTISAGIDGQPMRRSGDNTTYTDMRDNFYRAVROLTPGAK 120
DB 61 YCYFGRKGYMDIDDHATTISAGIDGQPMRRSGDNTTYTDMRDNFYRAVROLTPGAK 120
QY 121 RIGIEFDHVNLDPRQLEALPGVEFVDISQPSMMRTIKSLSEOKLIREGARVCDVGA 180
DB 121 RIGIEFDHVDLDFRRTLEALPGVEFVDISQPSMMRTIKSLSEOKLIREGARVCDVGA 180
QY 181 ACAAIKAGVBEHEVAIATTNAMIKEIKSPFVELMDTWTWFOGINTDGAHNPVTNRI 240
DB 181 ACVAAVKAGVBEHEVAIATTNAMIKEIKSPFVELMDTWTWFOGINTDGAHNPVTNRI 240
QY 241 VOSGDILSLNTPFMIFGYTTALERTLPCDHVDASLDITWEKNVAHRRGLEIKPGARCK 300
DB 241 VOSGDILSLNTPFMIFGYTTALERTLPCDHVDASLDITWEKNVAHRRGLEIKPGARCK 300
QY 301 DIAELNEMRYEMWDLKTRSPGYGHSFGVLCHYYGREAGVELREDIDTELKPGMYVSMEP 360
DB 301 DIAELNEMRYEMWDLKTRSPGYGHSFGVLCHYYGREAGVELREDIDTELKPGMYVSMEP 360
QY 361 MVMLEPGMAGAGYREHDLIVGEDGAENITGPPGPEHNIIRN 404
DB 361 MVMLEPGMAGAGYREHDLIVGEDGAENITGPPGPEHNIIRN 404

RESULT 6
US-10-251-078-12
Sequence 12, Application US/10251078
Publication No. US20030119084A1
GENERAL INFORMATION:
APPLICANT: Roche Diagnostics GmbH
APPLICANT: Shao, Zhixin
APPLICANT: Schmuck, Rainer
APPLICANT: Kratzsch, Peter
APPLICANT: Kenkies, Janet
APPLICANT: Weisner, Harald
TITLE OF INVENTION: Variants of an Erwinia-type creatinase
FILE REFERENCE: 20981 EP
CURRENT APPLICATION NUMBER: US/10/251,078
CURRENT FILING DATE: 2002-09-20
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PatentIn version 3.1
SEQ ID NO 12
LENGTH: 404
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:variant CT2m9
US-10-251-078-12

Query Match 94.2%; Score 2064; DB 14; Length 404;
Best Local Similarity 93.6%; Pred. No. 9.7e-194;
Matches 378; Conservative 16; Mismatches 10; Indels 0; Gaps 0;

QY 1 MTDDMLHMKHNGEKDYSPPSDAEMTRQNDVRGMKANNVDALFTSYHCINYSGL 60
DB 1 MTDDMLHMKHNGEKESPPSDAEMTRQSDVRMMAMENDVDALFTSYHCINYSGL 60
QY 61 YCYFGRKGYMDHNNATTISAGIDGQPMRRSGDNTTYTDMRDNFYRAVROLTPGAK 120
DB 61 YCYFGRKGYMDIDDHATTISAGIDGQPMRRSGDNTTYTDMRDNFYRAVROLTPGAK 120
QY 121 RIGIEFDHVNLDPRQLEALPGVEFVDISQPSMMRTIKSLSEOKLIREGARVCDVGA 180

Db 121 RIGIEFHDVLDLFRRTLEALPGVEFVDIGQPSMMKRVLSLEOKLIREGARICDVGA 180
Qy 181 ACAAIKAGVBEHEVALATTNAMIREIAKSPPEVEMDTWTWFOGINTDGAHPVTNRI 240
Db 181 ACVAAYKAGVBEHEVALATTNAVREIAKSPPEVEMDTWTWFOGINTDGAHPVTNRI 240
Qy 241 VQSGDIISLNTFPMIFGYTTLERTLFCDHVDASLDITWENKVAVHRGELIRPGARCK 300
Db 241 VQSGDIISLNTFPMIFGYTTLERTLFCDHVDASLDITWENKVAVHRGELIRPGARCK 300
Qy 301 DIAELNEMREMDLTKRSFGYGSFGVLCHYYGREGAVELREDIDTELKPGWVSWEP 360
Db 301 DIAELNEMREMDLTKRSFGYGSFGVLSHYGREGAVELREDIDTVLQPGWVSWEP 360
Qy 361 MVMLEPGMPGAGYREHDLIVGEDGAGNITGPPGPEHNITRN 404
Db 361 MVMLEPGMPGAGYREHDLIVGEDGAGNITGPPGPEHNITRN 404

RESULT 7
US-10-251-078-14
Sequence 14, Application US/10251078
Publication No. US20030119084A1
GENERAL INFORMATION:
APPLICANT: Roche Diagnostics GmbH
APPLICANT: Shao, Zhixin
APPLICANT: Schmuck, Rainer
APPLICANT: Kratzsch, Peter
APPLICANT: Kenkies, Janet
APPLICANT: Weisner, Harald
TITLE OF INVENTION: Variants of an Erwinia-type creatinase
FILE REFERENCE: 20981 EP
CURRENT APPLICATION NUMBER: US/10/251,078
CURRENT FILING DATE: 2002-09-20
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PatentIn version 3.1
SEQ ID NO 14
LENGTH: 404
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:variant CT2m10
US-10-251-078-14

Query Match 94.2%; Score 2062; DB 14; Length 404;
Best Local Similarity 93.6%; Pred. No. 1.5e-193;
Matches 378; Conservative 15; Mismatches 11; Indels 0; Gaps 0;
Qy 1 MTDDMLHVMKMHNGEKDYSFSDAEMTRRQNDVGMKAKNVDAALFTSHICINYSGL 60
Db 1 MTDDMLHVMKMHNGEKDYSFSDAEMTRRQSDVRRMWAENDVDALFTSHICINYSGL 60
Qy 61 YCYGRKKGWVIDDHNNATTISAGIDGQPMRRSFGDNITTYDMRRDNFYARVRLTGGAK 120
Db 61 YCYGRKKGWVIDDHNNATTISAGIDGQPMRRSFGDNITTYDMRRDNFYARVRLTGGAK 120
Qy 121 RIGIEFHDVLDLFRRTLEALPGVEFVDIGQPSMMKRVLSLEOKLIREGARICDVGA 180
Db 121 RIGIEFHDVLDLFRRTLEALPGVEFVDIGQPSMMKRVLSLEOKLIREGARICDVGA 180
Qy 181 ACAAIKAGVBEHEVALATTNAMIREIAKSPPEVEMDTWTWFOGINTDGAHPVTNRI 240
Db 181 ACVAAYKAGVBEHEVALATTNAVREIAKSPPEVEMDTWTWFOGINTDGAHPVTNRI 240
Qy 241 VQSGDIISLNTFPMIFGYTTLERTLFCDHVDASLDITWENKVAVHRGELIRPGARCK 300
Db 241 VQSGDIISLNTFPMIFGYTTLERTLFCDHVDASLDITWENKVAVHRGELIRPGARCK 300
Qy 301 DIAELNEMREMDLTKRSFGYGSFGVLCHYYGREGAVELREDIDTELKPGWVSWEP 360
Db 301 DIAELNEMREMDLTKRSFGYGSFGVLSHYGREGAVELREDIDTVLQPGWVSWEP 360

Qy 361 MVMLEPGMPGAGYREHDLIVGEDGAGNITGPPGPEHNITRN 404
Db 361 MVMLEPGMPGAGYREHDLIVGEDGAGNITGPPGPEHNITRN 404

RESULT 8
US-10-251-078-20
Sequence 20, Application US/10251078
Publication No. US20030119084A1
GENERAL INFORMATION:
APPLICANT: Roche Diagnostics GmbH
APPLICANT: Shao, Zhixin
APPLICANT: Schmuck, Rainer
APPLICANT: Kratzsch, Peter
APPLICANT: Kenkies, Janet
APPLICANT: Weisner, Harald
TITLE OF INVENTION: Variants of an Erwinia-type creatinase
FILE REFERENCE: 20981 EP
CURRENT APPLICATION NUMBER: US/10/251,078
CURRENT FILING DATE: 2002-09-20
NUMBER OF SEQ ID NOS: 32
SOFTWARE: PatentIn version 3.1
SEQ ID NO 20
LENGTH: 404
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:variant CTsd2
US-10-251-078-20

Query Match 94.0%; Score 2058; DB 14; Length 404;
Best Local Similarity 93.3%; Pred. No. 3.8e-193;
Matches 377; Conservative 16; Mismatches 11; Indels 0; Gaps 0;
Qy 1 MTDDMLHVMKMHNGEKDYSFSDAEMTRRQNDVGMKAKNVDAALFTSHICINYSGL 60
Db 1 MTDDMLHVMKMHNGEKDYSFSDAEMTRRQSDVRRMWAENDVDALFTSHICINYSGL 60
Qy 61 YCYGRKKGWVIDDHNNATTISAGIDGQPMRRSFGDNITTYDMRRDNFYARVRLTGGAK 120
Db 61 YCYGRKKGWVIDDHNNATTISAGIDGQPMRRSFGDNITTYDMRRDNFYARVRLTGGAK 120
Qy 121 RIGIEFHDVLDLFRRTLEALPGVEFVDIGQPSMMKRVLSLEOKLIREGARICDVGA 180
Db 121 RIGIEFHDVLDLFRRTLEALPGVEFVDIGQPSMMKRVLSLEOKLIREGARICDVGA 180
Qy 181 ACAAIKAGVBEHEVALATTNAMIREIAKSPPEVEMDTWTWFOGINTDGAHPVTNRI 240
Db 181 ACVAAYKAGVBEHEVALATTNAVREIAKSPPEVEMDTWTWFOGINTDGAHPVTNRI 240
Qy 241 VQSGDIISLNTFPMIFGYTTLERTLFCDHVDASLDITWENKVAVHRGELIRPGARCK 300
Db 241 VQSGDIISLNTFPMIFGYTTLERTLFCDHVDASLDITWENKVAVHRGELIRPGARCK 300
Qy 301 DIAELNEMREMDLTKRSFGYGSFGVLCHYYGREGAVELREDIDTELKPGWVSWEP 360
Db 301 DIAELNEMREMDLTKRSFGYGSFGVLSHYGREGAVELREDIDTVLQPGWVSWEP 360
Qy 361 MVMLEPGMPGAGYREHDLIVGEDGAGNITGPPGPEHNITRN 404
Db 361 MVMLEPGMPGAGYREHDLIVGEDGAGNITGPPGPEHNITRN 404

RESULT 9
US-10-369-493-21632
Sequence 21632, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng

;; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
;; FILE REFERENCE: 38-10(52052)B
;; CURRENT APPLICATION NUMBER: US/10/369,493
;; CURRENT FILING DATE: 2003-02-28
;; PRIOR APPLICATION NUMBER: US 60/360,039
;; PRIOR FILING DATE: 2002-02-21
;; NUMBER OF SEQ ID NOS: 47374
;; SEQ ID NO 21632
;; LENGTH: 351
;; TYPE: PRT
;; ORGANISM: Pyrococcus abyssi
US-10-369-493-21632

Query Match 11.3% Score 248.5; DB 15; Length 351;
Best Local Similarity 25.3%; Pred. No. 2e-15; Indels 41; Gaps 13;
Matches 92; Conservative 70; Mismatches 161

QY 36 WMAKNNVDALFTSYHCINYYSGMLYCYFGKRYGVNIDHNNATTISAGIDGQPMRRSFG 95
DB 14 FPNESIDAVLITKNPNYYLSGASPLAGY---LVNNDGATLVPELEYEMAKES-- 68
QY 96 DNTYTTDM-RDNFRAVRQLTTGAKRIGIEPDHNLDFRQLEALPGVEFVDSQSM 154
DB 69 -KIPVEKRRRDDFKVPE---GVKVGIR-GSLSYFVEDLKEKGISEKTKVDVIK 122
QY 155 WRTITKSLSEOKLIREGARVCDVGACAAIKAGVPEHVALATTNMIKIAKSPFV 214
DB 123 EMRVKSDSEETIKETAKCEIADKAVMAIEETBEKREIRAKVEYIMKONGAKKRPD 182
QY 215 ELMDTWTFQSGINTDGAHNPVTRIVOSGDLISLNTFPMFGYTTALERTLFCDHVDA 274
DB 183 -----TISAGYSALPHGVASDRIKIGDLVIDLALVNHVSDITRTVVGSPNEK 236
QY 275 SLDIWEKNVAHRRGLEIKRGARCKDAIENLMEYREMDLKTSPFGYSGSPVLCY 334
DB 237 QREIYEIYLEAKQAVEAARPGITTK---ELDSIAR--NIIR-EGYGYDF--IHS 285
QY 335 GREAGVELRE-----DIDTELKPGVVMSEPMVMLEPFGMPGAGGYEHDLIYGE 387
DB 286 GHGVGLEHMPGVSYQD-ETVLKEGMVITLEPGIYIK---FGGVRIEDTIVITKTGA 340
QY 388 ENIT 391
DB 341 RLRT 344

RESULT 10
US-09-815-242-10645
;; Sequence 10645, Application US/09815242
;; Patent No. US20020061569A1
;; GENERAL INFORMATION:
;; APPLICANT: Haselbeck, Robert
;; APPLICANT: Ohlsen, Karl L.
;; APPLICANT: Zykkind, Judith W.
;; APPLICANT: Wall, Daniel
;; APPLICANT: Trawick, John D.
;; APPLICANT: Carr, Grant J.
;; APPLICANT: Yamamoto, Robert T.
;; APPLICANT: Xu, H. Howard
;; TITLE OF INVENTION: Identification of Essential Genes in
;; TITLE OF INVENTION: Prokaryotes
;; FILE REFERENCE: ELITRA.011A
;; CURRENT APPLICATION NUMBER: US/09/815,242
;; CURRENT FILING DATE: 2001-03-21
;; PRIOR APPLICATION NUMBER: 60/191,078
;; PRIOR FILING DATE: 2000-03-21
;; PRIOR APPLICATION NUMBER: 60/206,848
;; PRIOR FILING DATE: 2000-05-23
;; PRIOR APPLICATION NUMBER: 60/207,727
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: 60/242,578
;; PRIOR FILING DATE: 2000-10-23

;; PRIOR APPLICATION NUMBER: 60/253,625
;; PRIOR FILING DATE: 2000-11-27
;; PRIOR APPLICATION NUMBER: 60/257,931
;; PRIOR FILING DATE: 2000-12-22
;; PRIOR APPLICATION NUMBER: 60/269,308
;; PRIOR FILING DATE: 2001-02-16
;; NUMBER OF SEQ ID NOS: 14110
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 10645
;; LENGTH: 354
;; TYPE: PRT
;; ORGANISM: Enterococcus faecalis
US-09-815-242-10645

Query Match 10.5% Score 229.5; DB 9; Length 354;
Best Local Similarity 24.7%; Pred. No. 1.5e-13;
Matches 98; Conservative 59; Mismatches 164; Indels 75; Gaps 14;

QY 26 MTRQNDVRGMMAKNNVDALFTSYHCINYYSGMLYCYFGKRYGVNIDHNNATTISAGID 85
DB 1 MAAVEKRLRELMKENNLQGLVTSPLYRLVLTN-----FTGTGLAM-----IAMD 46
QY 86 GGGPMRRSFGDNTYTTDMRDNFRAVRQLT--TGAKRIGIEPDH--NDF 133
DB 47 -----KAFITDFYTE-----QAAEQATGFTIVKNTGHIPEVDLAERLQDNLAF 94
QY 134 RRQ-----LEALPGVEFVDSQSMWRTITKSLSEOKLIREGARVCDVGACAA 184
DB 95 EETOVSPADYSLEIILP-CELVPMGLIELEREYKDEEVAIIIEKACALDQGAFLVE 153
QY 185 AIKAGVEHEVALATTNMIKIAKSPFVELMDTWTFQSGINTDGAHNPVTRIVOSG 244
DB 154 MIKGMTEIEVANQLDPMRSKASGVSEFTIV-----ASGLSAMPBGVASHVIRKG 207
QY 245 DILSLNTFPMFGYTTALERTLFCDHVDASLDIWEKNVAHRRGLEIKRGARCKDAI 304
DB 208 ELITLDRCGYEYGVSDMTRFPAIGSIQPKLKEIYDYLEQLVLAERKGL---TGI 263
QY 305 ELNMEYREMDLKTSPFGYSGSPVLCYHGREAGVELRE-----IDTELKPGVVM 358
DB 264 QLDALAND---HIASGYGDAFG---HSTHGIGLEIHEPNVSPFRADKQFVGNVITD 316
QY 359 EPMVMLEPFGMPGAGGYEHDLIYGEEDGAEINITGP 394
DB 317 EPGIYL-----PGIGVRIEDDLSTIAGNRVLTHAP 348

RESULT 11
US-09-815-242-10235
;; Sequence 10235, Application US/09815242
;; Patent No. US20020061569A1
;; GENERAL INFORMATION:
;; APPLICANT: Haselbeck, Robert
;; APPLICANT: Ohlsen, Karl L.
;; APPLICANT: Zykkind, Judith W.
;; APPLICANT: Wall, Daniel
;; APPLICANT: Trawick, John D.
;; APPLICANT: Carr, Grant J.
;; APPLICANT: Yamamoto, Robert T.
;; APPLICANT: Xu, H. Howard
;; TITLE OF INVENTION: Identification of Essential Genes in
;; TITLE OF INVENTION: Prokaryotes
;; FILE REFERENCE: ELITRA.011A
;; CURRENT APPLICATION NUMBER: US/09/815,242
;; CURRENT FILING DATE: 2001-03-21
;; PRIOR APPLICATION NUMBER: 60/191,078
;; PRIOR FILING DATE: 2000-03-21
;; PRIOR APPLICATION NUMBER: 60/206,848
;; PRIOR FILING DATE: 2000-05-23
;; PRIOR APPLICATION NUMBER: 60/207,727
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: 60/242,578
;; PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10235
LENGTH: 361
TYPE: PRT
ORGANISM: Escherichia coli
US-09-815-242-10235

Query Match 10.5%; Score 229; DB 9; Length 361;
Best Local Similarity 23.9%; Pred. No. 1.7e-13;
Matches 93; Conservative 64; Mismatches 162; Indels 70; Gaps 13;

QY 33 VRGMAKNNVDALFTSYH---CINYSGLYCYFGKKG-WYDHNATTISAGIDG 87
DB 7 LRDLKQQLDAVLISRKQKPHLGISTGSGYVVISRESAHLIVDSRYVEVEARAQGY 66
QY 88 QPWRSEFDNTYDWRDNFYRAVROLTTGAKRIGIEFDHYN-----LDFRQLE 139
DB 67 QHLHLDATNTLTIT---VNOIADDEQLQTL---IGFEGQVSMETAHRQSELNAKLVS 118
QY 140 ALPGVEFYDISPSMMRTIKSLBEOKLIRGARVCVGAACAALIKAGVPEHEVALAT 199
DB 119 ATPDV-----LRQITPREVEKIRLACGADRGAHEIRFIQAGSERIRIAEL 167
QY 200 TNAMIREAKSPFVELMDTWTWFOGSGINTGAGNPVTRIVQSGDILSLTFRIFERY 259
DB 168 EFWRQOGAERAKSEFDIV-----ASGWRGALPHGKADKIYAAEFVTLDFGALYQCYC 221
QY 260 TALERTLECD-----HVDASLDIWEKNVAVHRRGLIKPGARKDIAIENEMYR 311
DB 222 SDMTFTLLVNEGVSAESHL---LFTVQIVLQAGLAISAIRPVRCCQVDDAARVIT 278
QY 312 EMDLTKYRSPGYGHSFGVLCHYYGREGAVLELED-----IDTELKPMVVMSEPMWLP 365
DB 279 E-----AGYGVFG---HNTGALIGIEVHEDRFRSPRDTTLQPGMLTVERGIYL- 326
QY 366 EGMFGAGYRHHDLIYGEDGAENITGFP 394
DB 327 ---PGGGGVRIEDVVLVTPGABVLYAMP 352

RESULT 12

US-10-156-761-9699
Sequence 9699, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OMIURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHIRA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 9699
LENGTH: 381
TYPE: PRT
ORGANISM: Streptomyces avermitilis
US-10-156-761-9699

Query Match 9.3%; Score 203.5; DB 14; Length 381;
Best Local Similarity 24.3%; Pred. No. 5.8e-11;
Matches 97; Conservative 57; Mismatches 186; Indels 59; Gaps 15;

QY 14 GEXDYPSPDAEMTRQNDVRGMAKNNVDALFTSYHCINYSGLYCYFGKRYGMYL 72
DB 9 GTGDPAFTADYAAVRRAAGSAAADGDLVAPGPDLVWLGVPRVETRLTLVLR 68
QY 73 -DHNNATTI-----SAGIDGQFPWRSEFDNTYDWR---RDNFYAVROLTTGAKRIG 123
DB 69 AGHDPVLVPTLEAPDAABAAGP-----ALTRDMTGDKDP-YEAAATLDRSGRFG 120
QY 124 IEPDHNLDPRRQLEELPGVEFYDISPSMMRTIKSLBEOKLIRGARVCVGAACA 183
DB 121 IS-DNAMAMHLIGSGLPGTRYVALTEALPMLKAVDAABLE-----RIAAAGAAADA 173
QY 184 AAKR-----AGVEHEVAITTNAMIREIAKSPFVELMDTWTWFOGSGINTGAGNPVT 237
DB 174 TFEIRKVPFRAGRETDVA-----ADLAEILRQGHGQV---DFTVSGSPRGANPHHEAG 226
QY 238 NRIVQSGDILSLTFPMIFGYTALERTLPCDHVDDASLDIWEKNVAVHRRGLELIKPGA 297
DB 227 DVIATERGDMVLDGGLKHGYSPTSRTVHVGEBDEERKVDIVRAAOEAGFRAVRGV 286
QY 298 RCHDIAIELNEMREMLLKKRSFGYHSFGVLCHYYGREGAVLELED---IDTE---LK 351
DB 287 ACQDVDRABARVAD-----AGYDRP---IHRTHGIGIVTTHBPTWIEBORALV 335
QY 352 PGWVSMEMPMWLPBGMFGAGYRHHDLIYGEDGAENI 390
DB 336 PGMCFSEVPGVYL---PGRGVRIEDIVVTTDGGRL 370

RESULT 13

US-09-815-242-5287
Sequence 5287, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zykkind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: EILTRA, 011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5287
LENGTH: 350
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-5287

Query Match 9.1%; Score 200; DB 9; Length 350;
Best Local Similarity 21.7%; Pred. No. 1,1e-10;
Matches 83; Conservative 70; Mismatches 168; Indels 62; Gaps 12;

QY 29 RONDVGWAKNNVDAALFTSYHCINYYSGMLCYCFGRKGVNIDHNNATTISAGIDGQ 88
DB 1 RITGVHRIEKGKHLDAIILSDYNNRYLSC---FTGSGALILISKQYLIT----- 49
QY 89 PWRBSFGDNITTYTW-----RDNFYRAVRQL--TTGAKRIGIEFDHYNLDR 134
DB 50 -----DFRYIDATKQAPNVEIINRSTTIGIEKELHOENENVEFGSHVSYDXY 101
QY 135 RQLEBALPGVEFVDSQPSMMRRTIKSLBEQKLREGARVCVGAACAAIKAGVPEHE 194
DB 102 LELNKS--RISLISISNTVDKIRVDKADDELALLOKANIYDETEYIILTVKAGMTEGE 159
QY 195 VAIAATTNMIREIAKSPFVELMDTWTFQSGINTDGAHNPVTNRIVSGDILSLNTPPM 254
DB 160 LKAILESQKL-ELGADGSPFD-----TIVASGHRGALPHGVASDKIEKGDWITLDFGAY 213
QY 255 IFGYTALERTLFCHVDDASLDIWEKNVAHRRGLIELKPGARCKDIAIELNENYREND 314
DB 214 YNGYCSDIITRTFAIGEPDPKLEIYQIVLESQMKAINIIRPGMTGAE-ADALSRNYLE-- 270
QY 315 LKXRSFGYGRSFGYLCHYGRGAEVLE-----DIDTELKPGWVSMPEPMWLPBGM 368
DB 271 -----SKGYGKEFG---HSLHGIGLEIHEGPMALARTIODKLQVNNCTVBERGVYI----- 318
QY 369 PGAGYREHDILIVGEDGAENIT 391
DB 319 EGLGGIRIEDDILITENGCOVFT 341

Query Match 9.1%; Score 200; DB 9; Length 353;
Best Local Similarity 21.7%; Pred. No. 1,2e-10;
Matches 83; Conservative 70; Mismatches 168; Indels 62; Gaps 12;

QY 29 RONDVGWAKNNVDAALFTSYHCINYYSGMLCYCFGRKGVNIDHNNATTISAGIDGQ 88
DB 3 RITGVHRIEKGKHLDAIILSDYNNRYLSC---FTGSGALILISKQYLIT----- 51
QY 89 PWRBSFGDNITTYTW-----RDNFYRAVRQL--TTGAKRIGIEFDHYNLDR 134
DB 52 -----DFRYIDATKQAPNVEIINRSTTIGIEKELHOENENVEFGSHVSYDXY 103
QY 135 RQLEBALPGVEFVDSQPSMMRRTIKSLBEQKLREGARVCVGAACAAIKAGVPEHE 194
DB 104 LELNKS--RISLISISNTVDKIRVDKADDELALLOKANIYDETEYIILTVKAGMTEGE 161
QY 195 VAIAATTNMIREIAKSPFVELMDTWTFQSGINTDGAHNPVTNRIVSGDILSLNTPPM 254
DB 162 LKAILESQKL-ELGADGSPFD-----TIVASGHRGALPHGVASDKIEKGDWITLDFGAY 215
QY 255 IFGYTALERTLFCHVDDASLDIWEKNVAHRRGLIELKPGARCKDIAIELNENYREND 314
DB 216 YNGYCSDIITRTFAIGEPDPKLEIYQIVLESQMKAINIIRPGMTGAE-ADALSRNYLE-- 272
QY 315 LKXRSFGYGRSFGYLCHYGRGAEVLE-----DIDTELKPGWVSMPEPMWLPBGM 368
DB 273 -----SKGYGKEFG---HSLHGIGLEIHEGPMALARTIODKLQVNNCTVBERGVYI----- 320
QY 369 PGAGYREHDILIVGEDGAENIT 391
DB 321 EGLGGIRIEDDILITENGCOVFT 343

RESULT 14
US-09-815-242-12580
Sequence 12580, Application US/09815242
Patent No. US2002061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyekind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.01A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12580
LENGTH: 353
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-12580

RESULT 15
US-10-282-122A-70734
Sequence 70734, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl
APPLICANT: Zyekind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09

pb 241 TCGGCGGATCGAGCGGGCCAGCCCTGCGCGCGCAGCTTCGCGCAACATCACTAC 300
 Oy 301 ACCGACTGGCGCGCGGACATTTCTATGCGCGCGTGGCGCAGCTACCAACGGCGCCAG 360
 Db 301 ACCGACTGGCGCGCGGACATTTCTATGCGCGCGTGGCGCAGCTACCAACGGCGCCAG 360
 Oy 361 CGCATCGGATCGAGTTGACACGATGATCTGACTTCGCGCGCGCAGCTGAGGAACC 420
 Db 361 CGCATCGGATCGAGTTGACACGATGATCTGACTTCGCGCGCGCAGCTGAGGAACC 420
 Oy 421 CTACCGGCGCTGAGATTGCTGACATCAAGCCCTGATGTGATGCGACCATCAAG 480
 Db 421 CTACCGGCGCTGAGATTGCTGACATCAAGCCCTGATGTGATGCGACCATCAAG 480
 Oy 481 TCGCTGAAAGACAGAGCTGATCCGCGAAGGCGCGCTGTGTGAGTGTGCGCGCCG 540
 Db 481 TCGCTGAAAGACAGAGCTGATCCGCGAAGGCGCGCTGTGTGAGTGTGCGCGCCG 540
 Oy 541 GCGTGGCGGCGCTGACATCAAGGCGCGCGTGGCGCGAGATGAGTGGCGACCAACC 600
 Db 541 GCGTGGCGGCGCTGACATCAAGGCGCGCGTGGCGCGAGATGAGTGGCGACCAACC 600
 Oy 601 AATGCGATGATCCGCGAGATCGCAATGTTCCCTTCGTGAGCTGATGAGACCTGG 660
 Db 601 AATGCGATGATCCGCGAGATCGCAATGTTCCCTTCGTGAGCTGATGAGACCTGG 660
 Oy 661 ACTGCTTCAGTCCGCGCATCAACCGACGCGCGCAATCCGCTGACCAACCGCATC 720
 Db 661 ACTGCTTCAGTCCGCGCATCAACCGACGCGCGCAATCCGCTGACCAACCGCATC 720
 Oy 721 GTGCAATCGGCGGATCTTCTGCTCAACACCTTCCGATGATCTTGGGCTACTACG 780
 Db 721 GTGCAATCGGCGGATCTTCTGCTCAACACCTTCCGATGATCTTGGGCTACTACG 780
 Oy 781 GCGCTGAGCGCACGCTGTTCTGCGACCATGTGATGAGACGCCAGCTCGACACTGGAG 840
 Db 781 GCGCTGAGCGCACGCTGTTCTGCGACCATGTGATGAGACGCCAGCTCGACACTGGAG 840
 Oy 841 AAGAACTGGCGCTGATCGCGCGCGCTGAGCTGATCAAGCCGCGCGCGCTGCAAG 900
 Db 841 AAGAACTGGCGCTGATCGCGCGCGCTGAGCTGATCAAGCCGCGCGCGCTGCAAG 900
 Oy 901 GACATCGCATGAGCTCAACGAGATGTACCGCGAGTGGAGACTGCTGAGTACCGCTCC 960
 Db 901 GACATCGCATGAGCTCAACGAGATGTACCGCGAGTGGAGACTGCTGAGTACCGCTCC 960
 Oy 961 TTCCGCTATGCGCACTCTTCTGCGCTGCTGTGCTCACTACTAGGTGCGAGCGCGCTG 1020
 Db 961 TTCCGCTATGCGCACTCTTCTGCGCTGCTGTGCTCACTACTAGGTGCGAGCGCGCTG 1020
 Oy 1021 GAGCTGGCGAGACATGACACCGAGCTGAAAGCCCGGATGTGTCTTCCATGAGCCG 1080
 Db 1021 GAGCTGGCGAGACATGACACCGAGCTGAAAGCCCGGATGTGTCTTCCATGAGCCG 1080
 Oy 1081 ATGGGATGCTGCGGAGGAGATGCGCGGCTGCGGCGCTATCGGAGACGACATCTCG 1140
 Db 1081 ATGGGATGCTGCGGAGGAGATGCGCGGCTGCGGCGCTATCGGAGACGACATCTCG 1140
 Oy 1141 ATCGTGGGAGAGACGCTGCGGAGACATCAACGGCTTCCGTTGCTGCGGAGACAC 1200
 Db 1141 ATCGTGGGAGAGACGCTGCGGAGACATCAACGGCTTCCGTTGCTGCGGAGACAC 1200
 Oy 1201 ATCATCCGCAAC 1212
 Db 1201 ATCATCCGCAAC 1212

RESULT 2
 E12280 1212 bp DNA linear PART 27-APR-1998
 LOCUS E12280
 DEFINITION DNA encoding Alcaligenes thermostable creatin amidinohydrolase.
 ACCESSION E12280
 VERSION E12280.1 GI:3251114

KEYWORDS JP 1996308579-A/1.
 SOURCE Alcaligenes faecalis
 ORGANISM Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales; Alcaligenaceae; Alcaligenes.
 REFERENCE 1 (bases 1 to 1212)
 Sogabe,A., Yamamoto,K. and Kawamura,Y.
 GENE ENCODING CREATINE AMIDINOHYDROLASE
 Patent: JP 1996308579-A 1 26-NOV-1996;
 JOURNAL TOYOBO CO LTD
 COMMENT OS Alcaligenes faecalis
 PN JP 196308579-A/1
 PD 26-NOV-1996
 PF 16-MAY-1995 JP 1995117283
 PI SOGABE ATSUSHI, YAMAMOTO KAZUMI, KAWAMURA YOSHIHISA PC
 C12N15/09, C12N1/21, C12N9/78, (C12N15/09, C12R1:05), (C12N1/21, PC
 C12R1:425)
 PC (C12N9/78, C12R1:425);
 CC strandedness: Double;
 CC topology: Linear;
 CC hypothetical: No;
 CC anti-sense: No; Location/Qualifiers
 FH Key
 FT source 1..1212
 FT /organism='Alcaligenes faecalis' FT
 FT /clone='FE3581'
 FT mat_peptide 1..1212
 FT /product='Alcaligenes thermostable creatin FT
 FT amidinohydrolase'.

FEATURES
 source location/Qualifiers
 1..1212
 /organism='Alcaligenes faecalis'
 /mol_type='genomic DNA'
 /db_xref='taxon:511'

ORIGIN

Query Match 99.9%; Score 1210.4; DB 6; Length 1212;
 Best Local Similarity 99.9%; Pred. No. 5.7e-139;
 Matches 1211; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 ATGACTGACGACATGTTGACGATGATGAATGACCAACGCGGAGAAATATTTCGCG 60
 Db 1 ATGACTGACGACATGTTGACGATGATGAATGACCAACGCGGAGAAATATTTCGCG 60
 Oy 61 TTTTTCGATGCGGAGATGACCCCGCGCAAAACGATTCGCGGCTGAGTGGCAAGAC 120
 Db 61 TTTTTCGATGCGGAGATGACCCCGCGCAAAACGATTCGCGGCTGAGTGGCAAGAC 120
 Oy 121 AATGTCATGCGGCGCTGTTCACTTTATCACTGATCAACTACTATTTCGCGCTGGCTG 180
 Db 121 AATGTCATGCGGCGCTGTTCACTTTATCACTGATCAACTACTATTTCGCGCTGGCTG 180
 Oy 181 TACTGCTATTTCGAGCGCAAGTACGATGCTATGACCAACAAACGCCACGACATT 240
 Db 181 TACTGCTATTTCGAGCGCAAGTACGATGCTATGACCAACAAACGCCACGACATT 240
 Oy 241 TCGGCGGCGATCGACGCGCGCGACCCCTGGGCGCGAGCTTCGCGGCAACATCACTAC 300
 Db 241 TCGGCGGCGATCGACGCGCGCGACCCCTGGGCGCGAGCTTCGCGGCAACATCACTAC 300
 Oy 301 ACCGACTGGGCGCGCGCAATTTCTATCGCGCGTGGCGCGACATGACCAACGCGGCGCAAG 360
 Db 301 ACCGACTGGGCGCGCGCAATTTCTATCGCGCGTGGCGCGACATGACCAACGCGGCGCAAG 360
 Oy 361 CGCATCGGATCGAGTTGACACGATGATCTGACTTCGCGCGCGCAGCTGAGGAAGCC 420
 Db 361 CGCATCGGATCGAGTTGACACGATGATCTGACTTCGCGCGCGCAGCTGAGGAAGCC 420
 Oy 421 CTACCGGCGCTGAGTTGCTGACATCAAGCCCTGATGTGATGCGACCATCAAG 480
 Db 421 CTACCGGCGCTGAGTTGCTGACATCAAGCCCTGATGTGATGCGACCATCAAG 480

```

QY 481 TCGCTGAAGAGCAGAGCTGATCCGCGAAGCGCCGCGTGTGTGACGTCCGCGCGCG 540
DB 481 TCGCTGAAGAGCAGAGCTGATCCGCGAAGCGCCGCGTGTGTGACGTCCGCGCGCG 540
QY 541 GCTTCGCGCGCTGCATCAAGCGCGCGTGCCTCGAGCATGAAGTGGCGATCCAC 600
DB 541 GCTTCGCGCGCTGCATCAAGCGCGCGTGCCTCGAGCATGAAGTGGCGATCCAC 600
QY 601 AATGCGAGATCCGCGAGATCCGCGAATCGTTCCTTCGTGAGCTGATGGAACCTCG 660
DB 601 AATGCGAGATCCGCGAGATCCGCGAATCGTTCCTTCGTGAGCTGATGGAACCTCG 660
QY 661 ACCTGTTCCAGTCCGCGATCAACACGCGCGCGCAATCCGCGTCAACCAACCGCATC 720
DB 661 ACCTGTTCCAGTCCGCGATCAACACGCGCGCGCAATCCGCGTCAACCAACCGCATC 720
QY 721 GTGCAATCCGCGGAGATCTTTGCTCAACACCTTCCGATGATCTTGGCTACTACAC 780
DB 721 GTGCAATCCGCGGAGATCTTTGCTCAACACCTTCCGATGATCTTGGCTACTACAC 780
QY 781 GCGCTGAGCGCCAGCTGTTCTGCGACCAATGTCGATGACCGCAAGCTGACATCTGGAG 840
DB 781 GCGCTGAGCGCCAGCTGTTCTGCGACCAATGTCGATGACCGCAAGCTGACATCTGGAG 840
QY 841 AAGAAGCTGCGCGTGCATCGCCGCGGCTCGAGCTGATCAAGCCGCGCGCTGCAAG 900
DB 841 AAGAAGCTGCGCGTGCATCGCCGCGGCTCGAGCTGATCAAGCCGCGCGCTGCAAG 900
QY 901 GACATCGGCATCGAGCTCAACGAGATGTAACCGGAGTGGAGCTGCTGAAGTACCGCTCC 960
DB 901 GACATCGGCATCGAGCTCAACGAGATGTAACCGGAGTGGAGCTGCTGAAGTACCGCTCC 960
QY 961 TTCCGCTATGCGCACTCTTCCGCGCTGCTGCGCACTGTAAGCTGCGAGCGCGCGCTG 1020
DB 961 TTCCGCTATGCGCACTCTTCCGCGCTGCTGCGCACTGTAAGCTGCGAGCGCGCGCTG 1020
QY 1021 GAGCTGCGCGAGGACATCGACACCGAGCTGAAGCCGCGCATGATGATCTCCATGAGCGG 1080
DB 1021 GAGCTGCGCGAGGACATCGACACCGAGCTGAAGCCGCGCATGATGATCTCCATGAGCGG 1080
QY 1081 ATGCTGATGCTGCGCGAGGAGATGCGCGCGTGCCTGATCGGAGACGACATCTCTG 1140
DB 1081 ATGCTGATGCTGCGCGAGGAGATGCGCGCGTGCCTGATCGGAGACGACATCTCTG 1140
QY 1141 ATGCTGCGGAGAGACGCTGCGGAGACATCAACCGCGCTTCCGTTGCTCGGAAACAAAC 1200
DB 1141 ATGCTGCGGAGAGACGCTGCGGAGACATCAACCGCGCTTCCGTTGCTCGGAAACAAAC 1200
QY 1201 ATCATCCGCAAC 1212
DB 1201 ATCATCCGCAAC 1212

```

RESULT 3

```

LOCUS      E13584              1212 bp      DNA      linear      PAT 27-APR-1998
DEFINITION Alcaligenes faecalis gene for creatine amidinohydrolase.
ACCESSION  E13584
VERSION    E13584.1 GI:3252389
KEYWORDS   JP 1997215494-A/1.
SOURCE     Alcaligenes faecalis
ORGANISM   Alcaligenes faecalis
            Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
            Alcaligenaceae; Alcaligenes.
REFERENCE  1 (bases 1 to 1212)
            Sogabe,A., Hattori,T., Nishiyu,Y. and Kawamura,Y.
            NEW CREATINE AMIDINOHYDROLASE ITS PRODUCTION AND ITS USE
            JOURNAL Patent: JP 1997215494-A 1 19-AUG-1997;
            TOYOBO CO LTD
COMMENT    OS Alcaligenes faecalis
            PN JP 1997215494-A/1
            PD 19-AUG-1997
            PF 13-FEB-1996 JP 1996025435

```

```

PI SOGABE ATSUSHI, HATTORI TAKASHI, NISHIYA YOSHIKAKI, PI
KAWAMURA YOSHIKAKI
PC C12N9/78,C12N15/09,(C12N9/78,C12R1:05),(C12N15/09,C12R1:19);
CC strandedness: Double;
CC topology: linear;
FH Key Location/Qualifiers
FH source 1..1212
FT /strain='TB3581'
FT /organism='Alcaligenes faecalis' FT
FEATURES
source
1..1212
/organism='Alcaligenes faecalis'
/mol type='genomic DNA'
/db_xref='taxon:511'
ORIGIN
Query Match 99.9%; Score 1210.4; DB 6; Length 1212;
Best Local Similarity 99.9%; Pred. No. 5.7e-139;
Matches 1211; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 ATGACTGACGACATGTTGCAAGTGTGATAAATGCGACAAACGCGGAGAAAGATTATTCGCG 60
DB 1 ATGACTGACGACATGTTGCAAGTGTGATAAATGCGACAAACGCGGAGAAAGATTATTCGCG 60
QY 61 TTTTCGAGTCCGAGATGACCCGCGCCAAACGATTCGCGGCTGAGTGGCCAAAGAC 120
DB 61 TTTTCGAGTCCGAGATGACCCGCGCCAAACGATTCGCGGCTGAGTGGCCAAAGAC 120
QY 121 AATGTGATGCGCGCTGTTCACTCTTATGATGATCAACATAATTCGCGTGGCTG 180
DB 121 AATGTGATGCGCGCTGTTCACTCTTATGATGATCAACATAATTCGCGTGGCTG 180
QY 121 TACTGCTATTTCGAGCGCAATGAGGATGATGATGATGATGATGATGATGATGATGATGAT 240
DB 121 TACTGCTATTTCGAGCGCAATGAGGATGATGATGATGATGATGATGATGATGATGATGAT 240
QY 181 TACTGCTATTTCGAGCGCAATGAGGATGATGATGATGATGATGATGATGATGATGATGAT 240
DB 181 TACTGCTATTTCGAGCGCAATGAGGATGATGATGATGATGATGATGATGATGATGATGAT 240
QY 241 TCGGCGCGCATGACGCGCGCGAGCCCTGCGCGCGAGCTTCGCGGACAAATCACTAC 300
DB 241 TCGGCGCGCATGACGCGCGCGAGCCCTGCGCGCGAGCTTCGCGGACAAATCACTAC 300
QY 241 TCGGCGCGCATGACGCGCGCGAGCCCTGCGCGCGAGCTTCGCGGACAAATCACTAC 300
DB 241 TCGGCGCGCATGACGCGCGCGAGCCCTGCGCGCGAGCTTCGCGGACAAATCACTAC 300
QY 301 ACCGACTGCGCGCGCGAGCAATTTCTATGCGCGCGTGCAGCTGACCAACGCGCGCAAG 360
DB 301 ACCGACTGCGCGCGCGAGCAATTTCTATGCGCGCGTGCAGCTGACCAACGCGCGCAAG 360
QY 301 ACCGACTGCGCGCGCGAGCAATTTCTATGCGCGCGTGCAGCTGACCAACGCGCGCAAG 360
DB 301 ACCGACTGCGCGCGCGAGCAATTTCTATGCGCGCGTGCAGCTGACCAACGCGCGCAAG 360
QY 361 CGCATCGGCATGAGTTCGACACAGTCAATTCGATTCGCGCGCGAGCTGAGGAAAC 420
DB 361 CGCATCGGCATGAGTTCGACACAGTCAATTCGATTCGCGCGCGAGCTGAGGAAAC 420
QY 421 CTACCGGCGCTGAGTTCGTCGACATCAAGCCAGCCCTGATGATGATGATGATGATGATGAT 480
DB 421 CTACCGGCGCTGAGTTCGTCGACATCAAGCCAGCCCTGATGATGATGATGATGATGATGAT 480
QY 481 TCGCTGAAGAGCAGAGCTGATCCGCGAAGCGCGCGCTGCTGTGTGACGTGCGCGCGCG 540
DB 481 TCGCTGAAGAGCAGAGCTGATCCGCGAAGCGCGCGCTGCTGTGTGACGTGCGCGCGCG 540
QY 541 GCTTCGCGGCTGTCATCAAGCGCGCGTGCCTCGAGCATGAAGTGGCGATGCGCACAC 600
DB 541 GCTTCGCGGCTGTCATCAAGCGCGCGTGCCTCGAGCATGAAGTGGCGATGCGCACAC 600
QY 601 AATGCGAGTATCCGCGAGATCCGCGAATCGTTCCTTCGTGAGCTGATGGAACCTCG 660
DB 601 AATGCGAGTATCCGCGAGATCCGCGAATCGTTCCTTCGTGAGCTGATGGAACCTCG 660
QY 661 ACCTGTTCCAGTCCGCGATCAACACGCGCGCGCAATCCGCGTCAACCAACCGCATC 720
DB 661 ACCTGTTCCAGTCCGCGATCAACACGCGCGCGCAATCCGCGTCAACCAACCGCATC 720
QY 721 GTGCAATCCGCGGAGATCTTTGCTCAACACCTTCCGATGATCTTGGCTACTACAC 780
DB 721 GTGCAATCCGCGGAGATCTTTGCTCAACACCTTCCGATGATCTTGGCTACTACAC 780

```

QY 781 GGGCTGAGAGCGAGCTGTTCTGCGACATGTCGATGACGCCAGCTCGACATCTGGAG 840
 DB 781 GGGCTGAGAGCGAGCTGTTCTGCGACATGTCGATGACGCCAGCTCGACATCTGGAG 840
 QY 841 AAGAACGTGGCCGTGATGATCGCCGCGGCTCGAGCTGATCAAGCCGGCGCGCTGCAAG 900
 DB 841 AAGAACGTGGCCGTGATGATCGCCGCGGCTCGAGCTGATCAAGCCGGCGCGCTGCAAG 900
 QY 901 GACATCGCCATGAGCTCAACGAGATGATCCCGAGTGGAGACCTGTGTAAGTACCGCTCC 960
 DB 901 GACATCGCCATGAGCTCAACGAGATGATCCCGAGTGGAGACCTGTGTAAGTACCGCTCC 960
 QY 961 TTGGCTATGAGCACTCCCTTGGCGGCTGATGCACTACTACTAGGATGAGAGCGCGCGT 1020
 DB 961 TTGGCTATGAGCACTCCCTTGGCGGCTGATGCACTACTACTAGGATGAGAGCGCGCGT 1020
 QY 1021 GAGCTGCGGAGACATCGACACCGAGCTGAAGCCCGGAGTGTCTCCATGAGCGG 1080
 DB 1021 GAGCTGCGGAGACATCGACACCGAGCTGAAGCCCGGAGTGTCTCCATGAGAGCGG 1080
 QY 1081 ATGTGATGCTGCTCGGAGGCGATGCTCGGCTCGGCTATGCGAGACATCTCTG 1140
 DB 1081 ATGTGATGCTGCTCGGAGGCGATGCTCGGCTCGGCTATGCGAGACATCTCTG 1140
 QY 1141 ATGCTGCGGAGAGAGCGTCCGAGACATCACCGGCTTCCGTTGCTCGGAGACACAC 1200
 DB 1141 ATGCTGCGGAGAGAGCGTCCGAGACATCACCGGCTTCCGTTGCTCGGAGACACAC 1200
 QY 1201 ATCATCCGCAAC 1212
 DB 1201 ATCATCCGCAAC 1212

RESULT 4
 E16405 1212 bp DNA linear PART 28-JUN-1999
 LOCUS Alcaligenes faecalis gene for creatine amidinohydrolase.
 DEFINITION E16405.1 GI:5711088
 ACCESSION E16405.1 GI:5711088
 VERSION JP 1998174585-A/1.
 KEYWORDS Alkaligenes faecalis
 SOURCE Alkaligenes faecalis
 ORGANISM Alkaligenes faecalis
 Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 Alcaligenaceae; Alkaligenes.
 REFERENCE 1 (bases 1 to 1212)
 AUTHORS Sogabe, A., Nishiyu, Y., and Kawamura, Y.
 TITLE STABLE CREATINE AMIDINOHYDROLASE
 JOURNAL Patent: JP 1998174585-A 1 30-JUN-1998;
 TOYOBO CO LTD
 COMMENT OS Alkaligenes faecalis
 PN JP 1998174585-A/1
 PD 30-JUN-1998
 PI 17-DEC-1996 JP 1996337027
 PI SOGABE ATSUSHI, NISHIYU YOSHIAKI, KAWAMURA YOSHIOKA PC
 C12N9/78, C07H21/04, C12N1/21, C12N15/09//C12Q1/34, (C12N9/78, PC
 C12N1/19),
 PC (C12N1/21, C12N1/19);
 CC strandedness: Double;
 CC topology: Linear;
 FH Key Location/Qualifiers
 FT source 1. 1212
 FT mat_peptide 1. 1212
 FT /product='creatine amidinohydrolase'.
 FT Location/Qualifiers
 1. 1212
 /organism='Alkaligenes faecalis'
 /mol_type='genomic DNA'
 /db_xref='taxon:511'

FEATURES
 source
 1. 1212
 Location/Qualifiers

ORIGIN

Query Match 99.9%; Score 1210.4; DB 6; Length 1212;
 Best Local Similarity 99.9%; Pred. No. 5-7e-139;
 Matches 1211; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGACTGACGACATGTTGACAGTGAATGACCAACGCGGAGAAAGATTATTCGCG 60
 DB 1 ATGACTGACGACATGTTGACAGTGAATGACCAACGCGGAGAAAGATTATTCGCG 60
 QY 61 TTTTGGATGCGCGAGATGACCCGCGGCAAAAACGTTGCGGCTGATGAGCGCAAG 120
 DB 61 TTTTGGATGCGCGAGATGACCCGCGGCAAAAACGTTGCGGCTGATGAGCGCAAG 120
 QY 121 AATGCTGATGCGCGCTGTTCACTTATGACATGCACTACTACTATTCGCGCTG 180
 DB 121 AATGCTGATGCGCGCTGTTCACTTATGACATGCACTACTACTATTCGCGCTG 180
 QY 181 TACTGCTATTTTGGACGCGAAGTACGCGATGCTATGACACACACAAAGGCCAGCAT 240
 DB 181 TACTGCTATTTTGGACGCGAAGTACGCGATGCTATGACACACAAAGGCCAGCAT 240
 QY 241 TTGGCGGCGCATGACGCGGCGGCAAGCCGCGGCGGAGCTTCGCGACACATCACTAC 300
 DB 241 TTGGCGGCGCATGACGCGGCGGCAAGCCGCGGCGGAGCTTCGCGACACATCACTAC 300
 QY 241 TTGGCGGCGCATGACGCGGCGGCAAGCCGCGGCGGAGCTTCGCGACACATCACTAC 300
 DB 241 TTGGCGGCGCATGACGCGGCGGCAAGCCGCGGCGGAGCTTCGCGACACATCACTAC 300
 QY 301 ACCGACTGCGCGCGGCAATTTCTATGCGCGCGTGCACAGCTGACCAAGCGGCGCAAG 360
 DB 301 ACCGACTGCGCGCGGCAATTTCTATGCGCGCGTGCACAGCTGACCAAGCGGCGCAAG 360
 QY 361 CGCATGCGCATGAGATTGACACGTCATCTGCACTTCCGCGCGCACTGAGGAAGTC 420
 DB 361 CGCATGCGCATGAGATTGACACGTCATCTGCACTTCCGCGCGCACTGAGGAAGTC 420
 QY 421 CTACCGGCGCTGAGATTGTCGACATGACGCGGCTGATGATGATGATGATGATGATGAT 480
 DB 421 CTACCGGCGCTGAGATTGTCGACATGACGCGGCTGATGATGATGATGATGATGATGAT 480
 QY 481 TCGCTGAGAGACGAAAGCTGATCCGCGAAGCGCGCGTGTGTGAGCTGCGCGCGCG 540
 DB 481 TCGCTGAGAGACGAAAGCTGATCCGCGAAGCGCGCGTGTGTGAGCTGCGCGCGCG 540
 QY 541 GCTTGGCGGCTGCTGATCAAGCGCGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
 DB 541 GCTTGGCGGCTGCTGATCAAGCGCGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
 QY 601 AATGCTGATGATGCGGAGATGCGCAATGCTTCCCTGCGGAGCTGATGAGCACTTGG 660
 DB 601 AATGCTGATGATGCGGAGATGCGCAATGCTTCCCTGCGGAGCTGATGAGCACTTGG 660
 QY 661 ACCTGCTTCAAGTGGGATCAACACGAGCGCGGCAATTCGCTGACCAACCGCATC 720
 DB 661 ACCTGCTTCAAGTGGGATCAACACGAGCGCGGCAATTCGCTGACCAACCGCATC 720
 QY 721 GTGCAATCGCGGACATCTTTCGCTCAACACTTTCGCGATGATTTTGGCTACTACAC 780
 DB 721 GTGCAATCGCGGACATCTTTCGCTCAACACTTTCGCGATGATTTTGGCTACTACAC 780
 QY 781 GGGCTGAGAGCGAGCGTGTTCGCGACATGTCGATGAGCGGCTGACATCTGGAG 840
 DB 781 GGGCTGAGAGCGAGCGTGTTCGCGACATGTCGATGAGCGGCTGACATCTGGAG 840
 QY 841 AAGAACGTGGCCGTGATGATCGCCGCGGCTCGAGCTGATCAAGCCGGCGCGCTGCAAG 900
 DB 841 AAGAACGTGGCCGTGATGATCGCCGCGGCTCGAGCTGATCAAGCCGGCGCGCTGCAAG 900
 QY 901 GACATCGCCATGAGCTCAACGAGATGATCCCGAGTGGAGACCTGTGTAAGTACCGCTCC 960
 DB 901 GACATCGCCATGAGCTCAACGAGATGATCCCGAGTGGAGACCTGTGTAAGTACCGCTCC 960
 QY 961 TTGGCTATGAGCACTCCCTTGGCGGCTGATGCACTACTACTAGGATGAGAGCGCGT 1020
 DB 961 TTGGCTATGAGCACTCCCTTGGCGGCTGATGCACTACTACTAGGATGAGAGCGCGT 1020
 QY 1021 GAGCTGCGGAGACATCGACACCGAGCTGAAGCCCGGAGTGTCTCCATGAGAGCGG 1080

Db 1021 GAGCTGCGGAGAGACATCGACACCGAGCTGAAGCCCGGCGATGGTGTCTTCATGAGCCG 1080
Qy 1081 ATGTGATGCTCTCGGAGGGGCAATGCCCGTGCCTGCGGCTATCGCGAGAGACGATCTCG 1140
Db 1081 ATGTGATGCTCTCGGAGGGGCAATGCCCGTGCCTGCGGCTATCGCGAGAGACGATCTCG 1140
Qy 1141 ATGTGCGGAGAGAGAGCGTTCGAGACATCAACCGGCTTCCTGTCGGTCCGAGACACAAAC 1200
Db 1141 ATGTGCGGAGAGAGAGCGTTCGAGACATCAACCGGCTTCCTGTCGGTCCGAGACACAAAC 1200
Qy 1201 ATCATCCGCAAC 1212
Db 1201 ATCATCCGCAAC 1212
RESULT 5
BD017699
LOCUS BD017699 1212 bp DNA linear PAT 27-AUG-2002
DEFINITION Gene encoding creatine amidinohydrolase.
ACCESSION BD017699.1 GI:22558875
VERSION
KEYWORDS JP 2001252088-A/1.
SOURCE unclassified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 1212)
AUTHORS Sogabe,A., Yamamoto,K. and Kawamura,Y.
TITLE Gene encoding creatine amidinohydrolase
JOURNAL Patent: JP 2001252088-A 1 18-SEP-2001;
TOYOBO CO LTD
OS Alkaligenes faecalis TE3581 (FERM P14237)
PN JP 2001252088-A/1
PD 18-SEP-2001
PF 26-FEB-2001 JP 2001051054
PI ATSUSHI SOGABE, KAZUMI YAMAMOTO, YOSHIHISA KAWAMURA PC
C12N15/09, C12N1/21, C12N9/78//C12N15/09, C12R1:425), (C12N1/21, PC
C12R1:425),
PC (C12N9/78, C12R1:05), C12N15/00, (C12N15/00, C12R1:425) CC Gene
encoding creatine amidinohydrolase
FH Key Location/Qualifiers
FT source 1..1212
FT Location/Qualifiers
FT 1..1212
/organism='Alkaligenes faecalis TE3581 (FERM P14237)'.
/mol_type='genomic DNA'
/db_xref='taxon:32644'
FEATURES
source
ORIGIN
Query Match 99.9%; Score 1210.4; DB 6; Length 1212;
Best Local Similarity 99.9%; Pred. No. 5,7e-139;
Matches 1211; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ATGACTGCGACATGTTGACGTGATGAAATGGCAACGCGGAGAGAAATTAATTCGCG 60
Db 1 ATGACTGCGACATGTTGACGTGATGAAATGGCAACGCGGAGAGAAATTAATTCGCG 60
Qy 61 TTTTCGAGTGGCGAGATGACCGCGCCGCAAAACGACGTTTCGCGGCTGATGCGCAAGAC 120
Db 61 TTTTCGAGTGGCGAGATGACCGCGCCGCAAAACGACGTTTCGCGGCTGATGCGCAAGAC 120
Qy 121 AATGTGATGCGGCGCTGTTCACTCTTAATCACTGATCAACTACTATTCGCGCTGCTG 180
Db 121 AATGTGATGCGGCGCTGTTCACTCTTAATCACTGATCAACTACTATTCGCGCTGCTG 180
Qy 181 TACTGCTATTTCGAGCGAGATGAGCGGATGATCGACCAACGAGCGACGATTC 240
Db 181 TACTGCTATTTCGAGCGAGATGAGCGGATGATCGACCAACGAGCGACGATTC 240
Qy 241 TCGGCGGATCGACGCGCGGCGAGCCCTGCGGCGGCACTTCGCGGACAACTACCTAC 300
Db 241 TCGGCGGATCGACGCGCGGCGAGCCCTGCGGCGGCACTTCGCGGACAACTACCTAC 300

Db 241 TCGGCGGATCGACGCGCGGCGAGCCCTGCGGCGGCACTTCGCGGACAACTACCTAC 300
Qy 301 ACCGATCGGCGCGGAGCAATTTTATCGCGCGCGTGGCGCAGCTGACCAACGAGCGCGAAG 360
Db 301 ACCGATCGGCGCGGAGCAATTTTATCGCGCGCGTGGCGCAGCTGACCAACGAGCGCGAAG 360
Qy 361 CGCATCGGCAATCGATTGACACACGTCATCTGACCTTCGCGCGCAGCTCGAGAGAGCC 420
Db 361 CGCATCGGCAATCGATTGACACACGTCATCTGACCTTCGCGCGCAGCTCGAGAGAGCC 420
Qy 421 CTACCGGCGCTGAGTTGCTGACATCAAGCCAGCCCTTGATGATGATGCGACATCAAG 480
Db 421 CTACCGGCGCTGAGTTGCTGACATCAAGCCAGCCCTTGATGATGATGCGACATCAAG 480
Qy 481 TCGCTGAAAGAGACAGCTGATTCGGGAGAGCGCGCGGTGTGAGTGTGCGCGCGCG 540
Db 481 TCGCTGAAAGAGACAGCTGATTCGGGAGAGCGCGCGGTGTGAGTGTGCGCGCGCG 540
Qy 541 GCTTCGCGGCTGCCATCAAGCGCGCGTGCCTCGAGATGAAGTGGCGATCGCACACC 600
Db 541 GCTTCGCGGCTGCCATCAAGCGCGCGTGCCTCGAGATGAAGTGGCGATCGCACACC 600
Qy 601 AATGCGATGATCCGCGAGATGCGCAATCTGTTCCCTTCTGTGAGCTGATGACACTTGG 660
Db 601 AATGCGATGATCCGCGAGATGCGCAATCTGTTCCCTTCTGTGAGCTGATGACACTTGG 660
Qy 661 AACTGTTCGATGCGGCGATCAACACGAGCGCGCAATTCGCTGACCAACCGCATC 720
Db 661 AACTGTTCGATGCGGCGATCAACACGAGCGCGCAATTCGCTGACCAACCGCATC 720
Qy 721 GTGCAATCCGCGCATCTCTTCCTCAACACTTCCGATGATCTTCGCTACTACAC 780
Db 721 GTGCAATCCGCGCATCTCTTCCTCAACACTTCCGATGATCTTCGCTACTACAC 780
Qy 781 GCGCTGAGCGCAGCTGTTCTGCGACCATGTCATGACGCGACCTTGCATCTGGAG 840
Db 781 GCGCTGAGCGCAGCTGTTCTGCGACCATGTCATGACGCGACCTTGCATCTGGAG 840
Qy 841 AAGAACGTCGCGGATTCGCGCGGCGTGCAGCTGATCAAGCGCGCGCGCTGCAAG 900
Db 841 AAGAACGTCGCGGATTCGCGCGGCGTGCAGCTGATCAAGCGCGCGCGCTGCAAG 900
Qy 901 GACATCGCATCGAGCTCAACGAGATGATCCGCGAGTGGACCTGCTAAGTACCGCTCC 960
Db 901 GACATCGCATCGAGCTCAACGAGATGATCCGCGAGTGGACCTGCTAAGTACCGCTCC 960
Qy 961 TTGCGCTATGCGCACTCTTCGCGGCTGTGCACTAATCACTGCGGAGCGCGCTG 1020
Db 961 TTGCGCTATGCGCACTCTTCGCGGCTGTGCACTAATCACTGCGGAGCGCGCTG 1020
Qy 1021 GAGCTGCGGAGACATCGACACCGAGCTGAAGCCCGGATGCTTCCATGAGCGCG 1080
Db 1021 GAGCTGCGGAGACATCGACACCGAGCTGAAGCCCGGATGCTTCCATGAGCGCG 1080
Qy 1081 ATGTGATGCTGCGGAGAGGCGATGCGCGGTGCGGCGGCTATCGCGAGACGATCTCG 1140
Db 1081 ATGTGATGCTGCGGAGAGGCGATGCGCGGTGCGGCGGCTATCGCGAGACGATCTCG 1140
Qy 1141 ATGTGCGGAGAGAGAGCGTTCGAGACATCAACCGGCTTCCTGTCGGTCCGAGACAAAC 1200
Db 1141 ATGTGCGGAGAGAGAGCGTTCGAGACATCAACCGGCTTCCTGTCGGTCCGAGACAAAC 1200
Qy 1201 ATCATCCGCAAC 1212
Db 1201 ATCATCCGCAAC 1212
RESULT 6
BD105662
LOCUS BD105662 1212 bp DNA linear PAT 27-AUG-2002
DEFINITION Stable creatine amidinohydrolase.
ACCESSION BD105662
VERSION BD105662.1 GI:22651236

KEYWORDS JP 2001346594-A/1.
SOURCE unidentified
ORGANISM unclassified.

REFERENCE 1 (bases 1 to 1212)
AUTHORS Sogabe, A., Nishihara, Y., and Kawamura, Y.
TITLE Stable creatine amidinohydrolase
JOURNAL Patent: JP 2001346594-A 1 18-DEC-2001;
TOYOBO CO LTD
COMMENT OS Alcaligenes faecalis
PN JP 2001346594-A/1
PD 18-DEC-2001
PF 19-APR-2001 JP 2001121708
PI ATSUSHI SOGABE, YOSHIAKI NISHIYAMA, YOSHIOHISA KAWAMURA, PC
C12N15/00, C12N1/15, C12N1/19, C12N1/21, C12N5/10, C12N9/80// PC
C12N9/80, C12R1/05, C12N15/00, C12N5/00
CC Strandedness: Double;
CC Topology: Linear;
CC Stable creatine amidinohydrolase
FH Key Location/Qualifiers
FT source 1..1212 /organism='Alcaligenes faecalis',
location/Qualifiers
1..1212 /organism='unidentified',
/mol_type='genomic DNA',
/db_xref='taxon:32644'

ORIGIN

Query Match 99.9%; Score 1210.4; DB 6; Length 1212;
Best Local Similarity 99.9%; Pred. No. 5,7e-139;
Matches 1211; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGACTGACGACATGTTGCACTGATGAAATGGACAAACGCGAGAGAAATATTTCGCG 60
DB 1 ATGACTGACGACATGTTGCACTGATGAAATGGACAAACGCGAGAGAAATATTTCGCG 60
QY 61 TTTTTCGAGTGGCGAATGACCCGCGCCCAACGACGTTCCGCGGTGATGGCCCAAGAC 120
DB 61 TTTTTCGAGTGGCGAATGACCCGCGCCCAACGACGTTCCGCGGTGATGGCCCAAGAC 120
QY 121 AATGTCATGCGGCGCTGTTCACTTATCACTGATCACTAATCTATTTCGCGGTG 180
DB 121 AATGTCATGCGGCGCTGTTCACTTATCACTGATCACTAATCTATTTCGCGGTG 180
QY 181 TACTGCTATTTTCGACGCAAGTACGCAATGTCATGACCAACCAACGCGACGAT 240
DB 181 TACTGCTATTTTCGACGCAAGTACGCAATGTCATGACCAACCAACGCGACGAT 240
QY 241 TCGGCGCGCATGACGCGCGCCAGCCCTGCGCGCGCACTTCGGGACAACTACCTAC 300
DB 241 TCGGCGCGCATGACGCGCGCCAGCCCTGCGCGCGCACTTCGGGACAACTACCTAC 300
QY 241 TCGGCGCGCATGACGCGCGCCAGCCCTGCGCGCGCACTTCGGGACAACTACCTAC 300
DB 241 TCGGCGCGCATGACGCGCGCCAGCCCTGCGCGCGCACTTCGGGACAACTACCTAC 300
QY 301 ACCGACTGCGCGCGGACAACTTCTATGCGCGCGTGCAGCTGACCAACGCGCGCAAG 360
DB 301 ACCGACTGCGCGCGGACAACTTCTATGCGCGCGTGCAGCTGACCAACGCGCGCAAG 360
QY 361 CGCATGCGCATGACGATTCGACCACTGCACTTCGCGCGCGCACTTCGCGCGCAAG 420
DB 361 CGCATGCGCATGACGATTCGACCACTGCACTTCGCGCGCGCACTTCGCGCGCAAG 420
QY 421 CTACCGCGCGCTGACGATTCGACCACTGCACTTCGCGCGCGCACTTCGCGCGCAAG 480
DB 421 CTACCGCGCGCTGACGATTCGACCACTGCACTTCGCGCGCGCACTTCGCGCGCAAG 480
QY 481 TCGCTGAGAGACGACGATTCGCGCGAGCGCGCGCTGTGAGAGTGGCGCGCGCG 540
DB 481 TCGCTGAGAGACGACGATTCGCGCGAGCGCGCGCTGTGAGAGTGGCGCGCGCG 540
QY 541 GCGTGGCGCGCTGACGATTCGCGCGAGCGCGCGCTGTGAGAGTGGCGCGCGCG 600
DB 541 GCGTGGCGCGCTGACGATTCGCGCGAGCGCGCGCTGTGAGAGTGGCGCGCGCG 600

QY 601 AATGCGATGATCCGCGAGATGCGCAATGTTCCCTTCGTGAGCTGATGACACTG 660
DB 601 AATGCGATGATCCGCGAGATGCGCAATGTTCCCTTCGTGAGCTGATGACACTG 660
QY 661 ACTGTTTCCAGTGGCGATCAACACGAGCGCGCGCAATTCGGTTCACCAACGCGATC 720
DB 661 ACTGTTTCCAGTGGCGATCAACACGAGCGCGCGCAATTCGGTTCACCAACGCGATC 720
QY 721 GTGCAATCCGCGCGATCTTTCGCTCAACACTTCGCGATCTTCGCGCTACAC 780
DB 721 GTGCAATCCGCGCGATCTTTCGCTCAACACTTCGCGATCTTCGCGCTACAC 780
QY 781 GCGCTGAGCGACGCTGTTCTGCGACCATGTCGATGACGCGCACTTCGATCTGGAG 840
DB 781 GCGCTGAGCGACGCTGTTCTGCGACCATGTCGATGACGCGCACTTCGATCTGGAG 840
QY 841 AAGAACGTGGCGTGCATTCGCGCGCGCTGAGCTGATCAAGCGCGCGCGCTGCAG 900
DB 841 AAGAACGTGGCGTGCATTCGCGCGCGCTGAGCTGATCAAGCGCGCGCGCTGCAG 900
QY 901 GACATCCGCGCATGCGCTCAACGAGATGTAACCGCGAGTGGGACCTGTAAGTACCGCTCC 960
DB 901 GACATCCGCGCATGCGCTCAACGAGATGTAACCGCGAGTGGGACCTGTAAGTACCGCTCC 960
QY 961 TTCCGCTATGGCGCACTCTTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
DB 961 TTCCGCTATGGCGCACTCTTCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
QY 1021 GAGCTGCGCGAGACATGACACCGAGCTGAAAGCCGCGCATGCTGCTGCTGCTGCTG 1080
DB 1021 GAGCTGCGCGAGACATGACACCGAGCTGAAAGCCGCGCATGCTGCTGCTGCTGCTG 1080
QY 1081 ATGCTGATGCTGCGCGAGCGGATGCTCCGCGCGCGCTGCTGCTGCTGCTGCTGCTG 1140
DB 1081 ATGCTGATGCTGCGCGAGCGGATGCTCCGCGCGCGCTGCTGCTGCTGCTGCTGCTG 1140
QY 1141 ATGCTGCGGAGAGAGCGGTGCGGAGACATCAACGCGCTTCCGCTGCTGCTGCTGCTG 1200
DB 1141 ATGCTGCGGAGAGAGCGGTGCGGAGACATCAACGCGCTTCCGCTGCTGCTGCTGCTG 1200
QY 1201 ATCATCCGCAAC 1212
DB 1201 ATCATCCGCAAC 1212

RESULT 7
AB016788
LOCUS 1215 bp DNA linear BCT 06-JAN-2000
DEFINITION Alcaligenes sp. gene for creatine amidinohydrolase, complete cds,
strain:KS-85.
ACCESSION AB016788
VERSION AB016788.1 GI:6681665
KEYWORDS creatine amidinohydrolase.
SOURCE Alcaligenes sp.
ORGANISM Alcaligenes sp.
Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
Alcaligenaceae; Alcaligenes.
REFERENCE 1 (sites)
AUTHORS Furukawa, K., Ichikawa, T., Koyama, Y., and Suzuki, M.
TITLE Molecular cloning and sequence analysis of the gene encoding
thermostable creatinase from Alcaligenes sp. KS-85
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1215)
AUTHORS Koyama, Y. and Furukawa, K.
TITLE Direct Submission
JOURNAL Submitted (05-AUG-1998) Yasuji Koyama, Kikkoman Corporation,
Research and Development; Noda 399, Noda, Chiba 278-0037, Japan
(E-mail: dais-kim@92.so-net.ne.jp, Tel: +81-471-23-5571,
Fax: +81-471-23-5959)

FEATURES
source
1..1215
/organism='Alcaligenes sp.'
/mol_type='genomic DNA'

CDS

```

/strain="KS-85"
/db_xref="taxon:512"
1..1215
/codon_start=1
/translation="createine amidinohydrolyase"
/protein_id="BA88830.1"
/db_xref="GI:668166"
/translation="MTDDMLVMTKHNKEXDYPSPSDAEMTRBNDVYGMANXNVDA
ALFTSYHINYSGLCYRGRKGMVTDHNNATTSAGIDGGQWRSPDNTTYTD
WRDNFRAVRQLTTGAKRIGIEPDVNLDRRLDEALPGVEFDISQSPMNETIK
SLEBQKIRGARVCDVGAACAIAKGVBEHVALIPTNAMPRLKASPFVELMD
TWTFOSGINTDGAHPVTRNIVOSGDLISNTPEMIGYTALETILFCHVDLADL
DIKENVAVHRRGLKIKRGARCKDIAELNEMTREMDLKYRSFGYHRSFVLCYHY
GREGVELREDIDTEIKRGMVVSMEPMVMLBGMGAGVREHDLIVGEGDANITG
PPGPEHNITRN"

```

ORIGIN

```

Query Match      99.7%; Score 1208.8; DB 1; Length 1215;
Best Local Similarity 99.8%; Pred. No. 8.9e-139;
Matches 1210; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGACTGACGACATGTTGCAAGTGAATGACCAACGGCGAGAAATATTTCGCCG 60
DB 1 ATGACTGACGACATGTTGCAAGTGAATGACCAACGGCGAGAAATATTTCGCCG 60
QY 61 TTTTCGGATGCGCGATGACCCCGCGCAAAACGCTTCGCGCTGATGGCCAAAGAC 120
DB 61 TTTTCGGATGCGCGATGACCCCGCGCAAAACGCTTCGCGCTGATGGCCAAAGAC 120
QY 121 AATGTCGATGCGCGCTGTTCACTTATCACTGATCACTAATATTTCGGCTGCTG 180
DB 121 AATGTCGATGCGCGCTGTTCACTTATCACTGATCACTAATATTTCGGCTGCTG 180
QY 181 TACTGCTATTGTCGACGCAAGTACGGCATGTCATGACCAACAACGCCACGACGAT 240
DB 181 TACTGCTATTGTCGACGCAAGTACGGCATGTCATGACCAACAACGCCACGACGAT 240
QY 241 TCGGCGCGCATGACGCGCGCGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
DB 241 TCGGCGCGCATGACGCGCGCGCGACGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
QY 301 ACCGACTGCGCGCGCGCAATTTCTATGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
DB 301 ACCGACTGCGCGCGCGCAATTTCTATGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
QY 361 CGCATTCGGCATGAGTTGACCAAGTCAATCTGACCTTCGCGCGCGCGCGCGCG 420
DB 361 CGCATTCGGCATGAGTTGACCAAGTCAATCTGACCTTCGCGCGCGCGCGCGCG 420
QY 421 CTACCGCGCGCTGAGTTGTCGACATGACCAAGCTTGAATGATGATGCGCACTCA 480
DB 421 CTACCGCGCGCTGAGTTGTCGACATGACCAAGCTTGAATGATGATGCGCACTCA 480
QY 481 TCGCTCGAAGAGCAAGACTGATCCGCGAGCGCGCGCGCGCGCGCGCGCGCGCG 540
DB 481 TCGCTCGAAGAGCAAGACTGATCCGCGAGCGCGCGCGCGCGCGCGCGCGCGCG 540
QY 541 GCGTGCAGCGCTGTCATCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
DB 541 GCGTGCAGCGCTGTCATCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
QY 601 AATGCGATGATCCGCGAGATGCGCAATGCTTCCCTTCGTCGAGCTGATGACAC 660
DB 601 AATGCGATGATCCGCGAGATGCGCAATGCTTCCCTTCGTCGAGCTGATGACAC 660
QY 661 ACCTGATGATCCGCGAGATGCGCAATGCTTCCCTTCGTCGAGCTGATGACAC 720
DB 661 ACCTGATGATCCGCGAGATGCGCAATGCTTCCCTTCGTCGAGCTGATGACAC 720
QY 721 GTGCAATCCGCGACATCTTTCGCTCAACAACCTTCCGATGATCTTTCGCTACTAC 780
DB 721 GTGCAATCCGCGACATCTTTCGCTCAACAACCTTCCGATGATCTTTCGCTACTAC 780

```

```

QY 781 GCGCTGAGCGCGACGCTGTTCTGCGACCATGTCATGACGCGCGCGCGCGCGCG 840
DB 781 GCGCTGAGCGCGACGCTGTTCTGCGACCATGTCATGACGCGCGCGCGCGCGCG 840
QY 841 AAGAACGTCGCGTGCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
DB 841 AAGAACGTCGCGTGCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
QY 901 GACATCCGCAATCGACCTCAAGATGTAACGCGAGTGGACCTGCTGAAGTAC 960
DB 901 GACATCCGCAATCGACCTCAAGATGTAACGCGAGTGGACCTGCTGAAGTAC 960
QY 961 TTCGCTATGCGCACTCTCTTCGCGGCTGTCGCACTACTACGTCGCGAGCG 1020
DB 961 TTCGCTATGCGCACTCTCTTCGCGGCTGTCGCACTACTACGTCGCGAGCG 1020
QY 1021 GAGCTGCGCGAGACATGACACCGAGCTGAAGCCCGCGATGTCGTCATGAG 1080
DB 1021 GAGCTGCGCGAGACATGACACCGAGCTGAAGCCCGCGATGTCGTCATGAG 1080
QY 1081 ATGCTGATGCTGCGCGAGGCGATGCCGTCGCGCGCGCTATCGGACGACAT 1140
DB 1081 ATGCTGATGCTGCGCGAGGCGATGCCGTCGCGCGCGCTATCGGACGACAT 1140
QY 1141 ATCGTGGGAGAGACGCGTCCGAGACATCAACCGGCTTCCGTCGCGGAA 1200
DB 1141 ATCGTGGGAGAGACGCGTCCGAGACATCAACCGGCTTCCGTCGCGGAA 1200
QY 1201 ATCATCCGCAAC 1212
DB 1201 ATCATCCGCAAC 1212

```

```

RESULT 8
E11155
LOCUS E11155 1215 bp DNA linear PAT 29-SEP-1997
DEFINITION gDNA encoding creatine amidinohydrolyase.
ACCESSION E11155
VERSION E11155.1 GI:22024796
KEYWORDS UP 1996089255-A/1.
SOURCE unclassified
ORGANISM unclassified
REFERENCE 1 (bases 1 to 1215)
AUTHORS Furukawa,K., Ichikawa,T., Suzuki,M. and Koyama,T.
TITLE NOVEL CREATIVE AMIDINOHYDROLASE GENE, NOVEL RECOMBINANT DNA AND
JOURNAL PRODUCTION OF CREATIVE AMIDINOHYDROLASE
PATENT: JP 1996089255-A 1 09-APR-1996;
KIKKOMAN CORP
COMMENT OS Alcaligenes sp. KS-85
PN JP 1996089255-A/1
PD 09-APR-1996
PF 29-SEP-1994 JP 1994235737
PI FURUKAWA KEISUKE, ICHIKAWA TOSHIO, SUZUKI MASARU, KOYAMA TAJU
PC C12N15/09 C12N9/78, (C12N9/78, C12R1:05), (C12N9/78, C12R1:19); CC
strandedness: Double;
CC topology: linear;
FH Key Location/Qualifiers
FH source 1..1215 /organism="Alcaligenes sp. KS-85" FT CDS
FT 1..1215 /product="Creatine amidinohydrolyase".

```

FEATURES

```

source location/Qualifiers
1..1215 /organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

```

ORIGIN

```

Query Match      98.0%; Score 1188; DB 6; Length 1215;
Best Local Similarity 98.8%; Pred. No. 3.1e-136;

```

Matches 1197; Conservative 0; Mismatches 15; Indels 0; Gaps 0;	
Oy	1 ATGACTGACGACATTTTGCATGCTGATGAAATGCGCAACCGGGAAGAAATTAATTCGCGG 60
Db	1 ATGACTGACGACATTTTGCATGCTGATGAAATGCGCAACCGGGAAGAAATTAATTCGCGG 60
Oy	61 TTTTGGATGCGAGATGACCGCGCCAAAACGATTCGCGGCTGTGATGGCCAAAGAC 120
Db	61 TTTTGGATGCGAGATGACCGCGCCAAAACGATTCGCGGCTGTGATGGCCAAAGAC 120
Oy	121 AATGTCATGCGCGCTGCTTCACTTATATCATCTGATCACTACTATTCGGGCTGCTG 180
Db	121 AATGTCATGCGCGCTGCTTCACTTATATCATCTGATCACTACTATTCGGGCTGCTG 180
Oy	181 TACTGCTATTTTCGACGCAAGTACGGCATGTGATCGACCAACAAAGCCCAAGCATT 240
Db	181 TACTGCTATTTTCGACGCAAGTACGGCATGTGATCGACCAACAAAGCCCAAGCATT 240
Oy	241 TCGGCGGCGATCGACGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
Db	241 TCGGCGGCGATCGACGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
Oy	301 ACCGACTGCGCGCGGCAATTTCTATGCGCGCGCTGCGCAGCTGACCAAGCGGCGCAAG 360
Db	301 ACCGACTGCGCGCGGCAATTTCTATGCGCGCGCTGCGCAGCTGACCAAGCGGCGCAAG 360
Oy	361 CGCATCGCATGAGTTGACCAAGTCTCATCTTCCGCGGCGGCGGCGGCGGCGGCGGCGG 420
Db	361 CGCATCGCATGAGTTGACCAAGTCTCATCTTCCGCGGCGGCGGCGGCGGCGGCGGCGG 420
Oy	421 CTACCGGGCGTGAATGCTGACATCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480
Db	421 CTACCGGGCGTGAATGCTGACATCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480
Oy	481 TCGCTCGAAGAGCAGAACTGATTCGCGAAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540
Db	481 TCGCTCGAAGAGCAGAACTGATTCGCGAAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540
Oy	541 GCGTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600
Db	541 GCGTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600
Oy	601 AATGCGATGATCGCGGAGATCGCAAAATGCTTCCCTTGTGAGCTGATGAGCACTGCG 660
Db	601 AATGCGATGATCGCGGAGATCGCAAAATGCTTCCCTTGTGAGCTGATGAGCACTGCG 660
Oy	661 ACCTGCTTCACTGCGGCGATCAACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720
Db	661 ACCTGCTTCACTGCGGCGATCAACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720
Oy	721 GTGCAATCGGCGGCGATCTTTGCTCAACACTTTCGCGATGATCTTGGGCTACTACACC 780
Db	721 GTGCAATCGGCGGCGATCTTTGCTCAACACTTTCGCGATGATCTTGGGCTACTACACC 780
Oy	781 GCGCGTGAAGCGGCGGCTGTTCTGCGACATGTCATGATGAGCGGCGGCTGCACTTGGGAG 840
Db	781 GCGCGTGAAGCGGCGGCTGTTCTGCGACATGTCATGATGAGCGGCGGCTGCACTTGGGAG 840
Oy	841 AAGAACTGCGCGATCGCGCGGCGCTGAGCTGATCAAGCGGCGGCGGCGGCTGCGAG 900
Db	841 AAGAACTGCGCGATCGCGCGGCGCTGAGCTGATCAAGCGGCGGCGGCGGCTGCGAG 900
Oy	901 GACATCGCATGAGCTCAACGAGATGTCACCGGAGTGGAACTTGTGAAAGTACCGCTCC 960
Db	901 GACATCGCATGAGCTCAACGAGATGTCACCGGAGTGGAACTTGTGAAAGTACCGCTCC 960
Oy	961 TTTGCGCTAATGCGCATCTTTGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
Db	961 TTTGCGCTAATGCGCATCTTTGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
Oy	1021 GAGCTGCGGAGAGCATCGACCGAGCTGAAGCCCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080
Db	1021 GAGCTGCGGAGAGCATCGACCGAGCTGAAGCCCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080

Oy	1081 ATGGTATGCTGCGGAGGCGATCCCGGTCCCGGCTATCGGAGCAGACATCTCG 1140
Db	1081 ATGGTATGCTGCGGAGGCGATCCCGGTCCCGGCTATCCGAGCAGACATCTCG 1140
Oy	1141 ATGCTCGGAGGAGCGGTGCCGAGAACATCACCGGCTTCCGCTTCCGTCGCGAACAAC 1200
Db	1141 ATGCTCGGAGGAGCGGTGCCGAGAACATCACCGGCTTCCGCTTCCGTCGCGAACAAC 1200
Oy	1201 ATCATCGGCAAC 1212
Db	1201 ATCATCGGCAAC 1212
RESULT 9	
LOCUS	AX721933 1215 bp DNA linear PAT 07-MAY-2003
DEFINITION	Sequence 1 from Patent EP1298213.
ACCESSION	AX721933
VERSION	AX721933.1 GI:30422515
KEYWORDS	
SOURCE	Erwinia sp.
ORGANISM	Erwinia sp. Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales; Enterobacteriaceae; Erwinia.
REFERENCE	1 Shao, Z., Schmuck, R., Kratzsch, P., Kenkiles, J. and Weisner, H. Variants of an erwinia-type creatinase Patent: EP 1298213-A 1 02-APR-2003; Roche Diagnostics GmbH (DE); F. HOFMANN-LA ROCHE AG (CH)
FEATURES	location/Qualifiers 1..1215 /organism="Erwinia sp." /mol_type="unassigned DNA" /db_xref="taxon:558" 1..1215 /note="unnamed protein product" /codon_start=1 /transl_table=1 /protein_id="CAD90216.1" /db_xref="GI:30422515" /translation="MTDDMLHVMKANGKEKYSPSPDAEMTRROSDVPRMAENDVDA ALFTSYHCINYSGLFYCYFGRKYGWYIDQHAATTSAGIDGGQMRSPDNITTYTD WRDNFYQAVRLTPGARIGIEPFDHVLDPRLLEBALPVEFVDIGQPMRMSRPNMTVK SLEROKILREBARICDVGAACVAAGVSEHVAIATTNAYREIKSPFVELMD TWTWFOGINTDGAHNPVNRIVOSGILSNTPPMI FGYYTLERTLFCHVDASL DIWKRNVAVRHREGLIKPRARCDDIAIEMEMREMDLKYRSGVGSFVSHY GREAGVELRIDITVLQPGMVSMRPVMLPEGAPGAGYREHDLIVGBDAENITG PPFPEHNHIN"
CDS	
ORIGIN	
Query Match	81.3%; Score 984.8; DB 6; Length 1215;
Best Local Similarity	88.3%; Pred. No. 2.1e-111;
Matches 1070; Conservative 0; Mismatches 142; Indels 0; Gaps 0;	
Oy	1 ATGACTGACGACATTTTGCATGCTGATGAAATGCGCAACCGGGAAGAAATTAATTCGCGG 60
Db	1 ATGACTGACGACATTTTGCATGCTGATGAAATGCGCAACCGGGAAGAAATTAATTCGCGG 60
Oy	61 TTTTGGATGCGAGATGACCGCGCCAAAACGATTCGCGGCTGTGATGGCCAAAGAC 120
Db	61 TTTTGGATGCGAGATGACCGCGCCAAAACGATTCGCGGCTGTGATGGCCAAAGAC 120
Oy	121 AATGTCATGCGCGCTGCTTCACTTATATCATCTGATCACTACTATTCGGGCTGCTG 180
Db	121 AATGTCATGCGCGCTGCTTCACTTATATCATCTGATCACTACTATTCGGGCTGCTG 180
Oy	181 TACTGCTATTTTCGACGCAAGTACGGCATGTGATCGACCAACAAAGCCCAAGCATT 240
Db	181 TACTGCTATTTTCGACGCAAGTACGGCATGTGATCGACCAACAAAGCCCAAGCATT 240
Oy	241 TCGGCGGCGATCGACGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
Db	241 TCGGCGGCGATCGACGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300

Db 541 GCCCTGCTCCCGCCGTCAGAGCCGCGCTTCCGAGACGAGGTCGCGATCCGACACACC 600
 QY 601 AATGCAGTATCCGCGAGATGCGCAATGGTTCCTCCCTGCTGAGGCTATGACACCTGG 660
 Db 601 AATGCAGTATCCGCGAGATGCGCAATGGTTCCTCCCTGCTGAGGCTATGACACCTGG 660
 QY 661 ACCCTGTTCCAGTCGCGGATCAACACGAGCGCGGACATCCGCTGACCAACCGCATC 720
 Db 661 ACCCTGTTCCAGTCGCGGATCAACACGAGCGCGGACATCCGCTGACCAACCGCATC 720
 QY 721 GTGCATATCCGCGGACATCTCTTTCGCTCAACACTTCCCGCATATCTTCGCTCAAC 780
 Db 721 GTGCATATCCGCGGACATCTCTTTCGCTCAACACTTCCCGCATATCTTCGCTCAAC 780
 QY 781 GCGCTGAGAGCGACGCTGTTCTGCGACCATGTGATGAGCGCGACGCTCGACATCTGAG 840
 Db 781 GCGCTGAGAGCGACGCTGTTCTGCGACCATGTGATGAGCGCGACGCTCGACATCTGAG 840
 QY 841 AAGAACTGCGCGTGCATGCGCCGCGGAGCTCGAGCTGATCAAGCCGCGCGCGCTGCA 900
 Db 841 AAGAACTGCGCGTGCATGCGCCGCGGAGCTCGAGCTGATCAAGCCGCGCGCGCTGCA 900
 QY 901 GACATGCGCATCGAGCTCAACGAGATGACCGGAGTGGAGTCTGCTGAAGTACCGCTCC 960
 Db 901 GATATGCGCATCGAGCTCAACGAGATGACCGGAGTGGAGTCTGCTGAAGTACCGCTCC 960
 QY 961 TTCCGCTATGCGCATCTTCCGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
 Db 961 TTCCGCTATGCGCATCTTCCGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
 QY 1021 GAGCTGCGGAGGACATCGACACCGAGCTGAGAGCCGCGGACATGCTGCTGCTGAGCCG 1080
 Db 1021 GAGCTGCGGAGGACATCGACACCGAGCTGAGAGCCGCGGACATGCTGCTGCTGAGCCG 1080
 QY 1081 ATGCTGATCTGCTGCGGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
 Db 1081 ATGCTGATCTGCTGCGGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
 QY 1141 ATGCTGCGGAGGAGCGGTCGCGGAGACATCAACGCGCTTCCGCTGCTGCTGCGAACA 1200
 Db 1141 ATGCTGCGGAGGAGCGGTCGCGGAGACATCAACGCGCTTCCGCTGCTGCTGCGAACA 1200
 QY 1201 ATCATCCGCAAC 1212
 Db 1201 ATCATCCGCAAC 1212

RESULT 11
 AX721941 1215 bp DNA linear PART 07-MAY-2003
 LOCUS AX721941 Sequence 9 from Patent BP1298213.
 DEFINITION AX721941
 ACCESSION AX721941 GI:30422517
 VERSION AX721941.1
 KEYWORDS
 SOURCE
 ORGANISM
 synthetic construct
 synthetic construct
 other sequences; artificial sequences.
 REFERENCE
 1 Shao, Z., Schumuck, R., Kratzsch, P., Kenklies, J. and Weisner, H.
 TITLE Variants of an erwinia-type creatinase
 JOURNAL Patent: EP 1298213-A 9 02-APR-2003;
 Roche Diagnostics GmbH (DE) ; F.HOFFMANN-LA ROCHE AG (CH)
 FEATURES
 location/Qualifiers
 1..1215
 /organism="synthetic construct"
 /mol_type="unassigned DNA"
 /db_xref="taxon:32630"
 /note="variant CT1m24"
 1..1215
 /note="unnamed protein product"
 /codon_start=1
 /transl_table=11
 /protein_id="CAD90217.1"

ORIGIN

Query Match 81.0%; Score 981.6; DB 6; Length 1215;
 Best Local Similarity 88.1%; Pred. No. 5.2e-111;
 Matches 1068; Conservative 0; Mismatches 144; Indels 0; Gaps 0;

/db_xref="GI:30422518"
 /translation="MTDDMLHYMKHNEKEYSPEPSDAEMTRROSVDREMAENDVDA
 ALFTSYHCINYSGLFYCFGRKGMVLDIDDAATISAGIDGQPMRSGFNITFTD
 WRDNPYQAVNQLTPRGARRIGIEPHVDLDEPRTEBALRGYEPVDIGQSPMRRVTK
 SLEBCKLIREGRKIDVAGACVAAVKAGVPEHEVALATNNAWREBIASFPVEIAMD
 TWTFQSGINTDGANPVTNRIVQSGDILSTPFTFQYVTALEETLDCDHVDSL
 DTWEQNVAVHRRGELIKRGACKDIAELEMYEMBDLTKTRSPGYGHSFGLSHY
 GSEAVELREDDIDVTIQPMVMSBPMPLPEGAPGAGGVRSHDILLIYVDEDAENITG
 PFGPEHNIIRN"

QY 1 ATGACTGACGAGATGATGACGTGATGAAATGGCAACCGGAGAAAGATTATTCGCGG 60
 Db 1 ATGACTGACGAGATGATGACGTGATGAAATGGCAACCGGAGAAAGATTATTCGCGG 60
 QY 61 TTTTGGATGCGGAGATGACCGCGCCCAAAACGATTCGCGGCTGATGCGCAAGAAC 120
 Db 61 TTTTGGATGCGGAGATGACCGCGCCCAAAACGATTCGCGGCTGATGCGCAAGAAC 120
 QY 121 AATGTCGATGCGGCGCTGTTCACTTCTTATCACTGCAATCACTATTCGCGCTG 180
 Db 121 GACGTGACGCTGCGCTGTTCACTTCTTATCACTGCAATCACTATTCGCGCTG 180
 QY 181 TACTGCTATTTTGGACGCAAGTACGCGCATGTCATGACCAACAAAGCCCAAGCAT 240
 Db 181 TACTGCTATTTTGGACGCAAGTACGCGCATGTCATGACCAAGCCCAAGCAT 240
 QY 241 TCGGCGGCGATGACGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
 Db 241 TCGGCGGCGATGACGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
 QY 301 ACCGACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
 Db 301 ACCGACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
 QY 361 CGGATCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
 Db 361 CGGATCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
 QY 421 CTACCGGCGCGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
 Db 421 CTACCGGCGCGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
 QY 481 TCGCTGAAAGACAGACCTGATCGCGAAGGCGCGCGCTGTGATGATGATGATGATGAT 540
 Db 481 TCGCTGAAAGACAGACCTGATCGCGAAGGCGCGCGCTGTGATGATGATGATGATGAT 540
 QY 541 GCGTTCGAGGCGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
 Db 541 GCGTTCGAGGCGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
 QY 601 AATGCAGTATCCGCGGAGATGCGCAATGCTTCCCTTCGTGAGAGCTGATGAGACCTGG 660
 Db 601 AATGCAGTATCCGCGGAGATGCGCAATGCTTCCCTTCGTGAGAGCTGATGAGACCTGG 660
 QY 661 ACCCTGTTCCAGTCGCGGATCAACACGAGCGCGGCAATCCGCTGACCAACCGCATC 720
 Db 661 ACCCTGTTCCAGTCGCGGATCAACACGAGCGCGGCAATCCGCTGACCAACCGCATC 720
 QY 721 GTGCATATCCGCGGACATCTCTTTCGCTCAACACTTCCCGCATATCTTCGCTCAAC 780
 Db 721 GTGCATATCCGCGGACATCTCTTTCGCTCAACACTTCCCGCATATCTTCGCTCAAC 780
 QY 781 GCGCTGAGAGCGACGCTGTTCTGCGACCATGTGATGAGCGCGACGCTCGACATCTGAG 840
 Db 781 GCGCTGAGAGCGACGCTGTTCTGCGACCATGTGATGAGCGCGACGCTCGACATCTGAG 840
 QY 841 AAGAACTGCGCGTGCATGCGCCGCGGAGCTCGAGCTGATCAAGCCGCGCGCTGCAAG 900
 Db 841 AAGAACTGCGCGTGCATGCGCCGCGGAGCTCGAGCTGATCAAGCCGCGCGCTGCAAG 900

Db 841 AAGACGTGCGCGTGCACCGCGCGGCTCGAACTCATCAAGCCGGGTGCCCGCTGCAG 900
Qy 901 GAAATGCGCATTCAGCTGCACGAGATGTACCGGAGAGGAGACCTGCTGAATACCGCTCC 960
Db 901 GAAATGCGCATTCAGCTGCACGAGATGTACCGGAGAGGAGATCTGCTGAATACCGCTCC 960
Qy 961 TTGCGGTATGCGCATCTCTTCCGCGTGTGCTGCACTACTACCGTCCGAGGCGCGGTG 1020
Db 961 TTGCGGTATGCGCATCTCTTCCGCGTGTGCTGCACTACTACCGTCCGAGGCGCGGTG 1020
Qy 1021 GAGCTGCGGAGGACATCGACACCGAGCTGAAGCCCGCATGTGTGTCTTCATGAGCCG 1080
Db 1021 GAACTGCGGAGGACATCGATCCGTGCTGACCGCGCATGTGTGTCTTCATGAGCCG 1080
Qy 1081 ATGCTATGCTGCGGAGGAGCATGCGCGGTGCGCGCGCTATGCGGAGCAAGCATCTG 1140
Db 1081 ATGCTATGCTGCGGAGGAGCATGCGCGGTGCGCGCGCTATGCGGAGCAAGCATCTG 1140
Qy 1141 ATGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200
Db 1141 ATGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200
Qy 1201 ATCATCTCGCAAC 1212
Db 1201 ATCATCTCGCAAC 1212

RESULT 12
AX721943 1212 bp DNA linear PAT 07-MAY-2003
LOCUS AX721943
DEFINITION Sequence 11 from Patent EP1298213.
ACCESSION AX721943
VERSION AX721943.1 GI:30422519
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1 other sequences; artificial sequences.
AUTHORS Shao, Z., Schmuck, R., Kratzsch, P., Kenkies, J. and Weisser, H.
TITLES Variants of an erwinia-type creatinase
JOURNAL Patent: EP 1298213-A 11 02-APR-2003;
Roche Diagnostics GmbH (DE) ; F.HOFFMANN-LA ROCHE AG (CH)
LOCATION/Qualifiers
FEATURES
source 1..1212
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="variant CT2m9"
1..>1212
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAD90218.1"
/db_xref="GI:30422520"
/translation="MTDDMLHWKMHNGEKEYSPFSDAENTRQSDVRMAMENDVDA
ALPTSHCHINYSGRFLCYTGKRGKAVIIODHATTISAGIDGQDPRBSRGNTITTD
WRDNFTQAVROLTPGARRTGIFDHVDLFRITLSEALPGFVIVIGCPSPMRATVK
SLEBQULIREGARIQVGVACVAAVKAGVEHEVALATTNAVERIASFPVELDSL
TWTWPGSINTIDANHPVNRIVOSGDIISLNTFPMIFGYTALBERTLFCDHVDLSL
DTWEKVVAVHRRLLEIKPGARCKDIALIENEMYREMDLTKYSPFGSHFVLSHY
GRAGVELREDIDTIVLQPGWVSMEMPMVLPAGAPAGGRHMDILIVGDAENITG
PFGPEHNTIRN"

ORIGIN
Query Match 80.9%; Score 980; DB 6; Length 1212;
Best Local Similarity 88.0%; Pred. No. 8.2e-111;
Matches 1067; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

Qy 1 ATGACTGACGACATGTTGCACTGATGAATGGCAACGCGAAGAAATTTTGGCCG 60
Db 1 ATGACTGACGACATGTTGCACTGATGAATGGCAACGCGAAGAAATTTTGGCCG 60
Qy 61 TTTTGGAGTCCGAGATGACCGCGCGCAAAAGAGCTTGGCGGCTGAGTGGCAAGAC 120

Db 61 TTTTGGAGTCCGAGATGACCGCGCGCAAAAGAGCTTGGCGGCTGAGTGGCAAGAC 120
Qy 121 AATGTGAGATGCGCGCGCTGTTTCACTCTTAATCACTGATCAATCTACTTTCGGCTG 180
Db 121 GAGGTGAGCGCTGCGCTGTTTCACTCTTAATCACTGATCAATCTACTTTCGGCTG 180
Qy 181 TACTGCTATTTGCGAGCAAGTACGAGCATGTGATGACCAACAACGCGCAGAGATT 240
Db 181 TACTGCTATTTGCGAGCAAGTACGAGCATGTGATGACCAACAACGCGCAGAGATT 240
Qy 241 TCGGCGGAGTGCACGCGCGCGAGCCCTGCGCGCGCACTTCCGCGCAAACTACCTAC 300
Db 241 TCGGCGGAGTGCATGCGCGCGCGAGCCCTGCGCGCGCACTTCCGCGCAAACTACCTAT 300
Qy 301 ACCGATGCGCGCGCGCAAAATTTCTATTCGCGCGCGCGCGCACTTCCGCGCAAG 360
Db 301 ACCGATGCGCGCGCGCGCAAAATTTCTATTCGCGCGCGCGCGCACTTCCGCGCAAG 360
Qy 361 CGCATGCGCATGAGTTTGCACCACTGCACTTCCGCGCGCGCGCACTTCCGCGCAAG 420
Db 361 CGCATGCGCATGAGTTTGCACCACTGCACTTCCGCGCGCGCGCACTTCCGCGCAAG 420
Qy 421 CTACCGGCGCGTGCAGTTTGCACCACTGCACTTCCGCGCGCGCGCACTTCCGCGCAAG 480
Db 421 CTACCGGCGCGTGCAGTTTGCACCACTGCACTTCCGCGCGCGCGCACTTCCGCGCAAG 480
Qy 481 TCGCTGAGAGAGACGATGATCCGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
Db 481 TCGCTGAGAGAGACGATGATCCGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
Qy 541 GCGTGGCGCGCTGCGCATCAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
Db 541 GCGTGGCGCGCTGCGCATCAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
Qy 601 AATGCGATGATCCGCGAGATCCGCAATGCTTCCCTTCTGAGCTGATGACACTGG 660
Db 601 AATGCGATGATCCGCGAGATCCGCAATGCTTCCCTTCTGAGCTGATGACACTGG 660
Qy 661 ACTGCTTCCATGCTGCGGATCAACACCGAGCGCGCGCGCGCGCGCGCGCGCGCG 720
Db 661 ACTGCTTCCATGCTGCGGATCAACACCGAGCGCGCGCGCGCGCGCGCGCGCGCG 720
Qy 721 GTGCAATCGGCGGACATCTCTTCTGCGCAACCTTCCGATGATCTTCCGCTACTAC 780
Db 721 GTGCAATCGGCGGATCTCTTCTGCGCAACCTTCCGATGATCTTCCGCTACTAC 780
Qy 781 GCGCTGAGACGACGCTGTTTCTGCGACATGCTGATGAGCGCGCGCGCGCGCGCG 840
Db 781 GCGCTGAGACGACGCTGTTTCTGCGACATGCTGATGAGCGCGCGCGCGCGCGCG 840
Qy 841 AAGACGTGCGCGTGCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
Db 841 AAGACGTGCGCGTGCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
Qy 901 GACATGCGCATGAGTGCACGAGATGTACCGGAGTGGGACCTGTGTAAGTACCGCTC 960
Db 901 GACATGCGCATGAGTGCACGAGATGTACCGGAGTGGGATGCTGTGAAGTACCGCTC 960
Qy 961 TTGCGGTATGCGCATCTCTTCCGCGTGTGCTGCACTACTACCGTCCGAGGCGCGGTG 1020
Db 961 TTGCGGTATGCGCATCTCTTCCGCGTGTGCTGCACTACTACCGTCCGAGGCGCGGTG 1020
Qy 1021 GAGCTGCGGAGGACATCGACACCGAGCTGAAGCCCGCATGTGTGTCTTCATGAGCCG 1080
Db 1021 GAACTGCGGAGGACATCGATCCGTGCTGACCGCGCATGTGTGTCTTCATGAGCCG 1080
Qy 1081 ATGCTATGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
Db 1081 ATGCTATGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
Qy 1141 ATGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200

Db 1141 ATGCTGGGGAGACGCGCGCAGAGAAATTAACCGATTCCCTTGGCGGCTGAGACAAAC 1200
QY 1201 ATCATCCGCAAC 1212
Db 1201 ATCATCCGCAAC 1212
RESULT 13
AX721945 1212 bp DNA linear PAT 07-MAY-2003
LOCUS Sequence 13 from Patent EP1298213.
DEFINITION AX721945
ACCESSION AX721945
VERSION AX721945.1 GI:30422521
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE 1
AUTHORS Shao,Z., Schmuck,R., Kratzsch,P., Kenlles,J. and Weisner,H.
TITLE Variants of an erwinia-type creatinase
JOURNAL Patent: EP 1298213-A 13 02-APR-2003;
Roche Diagnostics GmbH (DE) ; F.HOFFMANN-LA ROCHE AG (CH)
FEATURES
source 1. 1212
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="variant CT2m10"
1. >1212
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAD90219.1"
/db_xref="GI:30422522"
/translation="MTDMLHYMKHNGEKEYSPFSDAMTRROSQVRRMAENDVDA
ALPTSHCHINYSGLPYCYFGRKGVVDDODHATTISAGIDGQPMRBSFGNITYTD
WRDNPYQAVROULTFGARIGIEPHVVDIPRTTLEBALPGEPIDGQSPMMWETVK
SIEEKLIREGARICDVGAGACVAAYKAVGPHEVALATINWVEIASFPFVELMD
TWTWQSGINTGAGHNPVTRNRIYVSGDILSLNTFPMI FGYYALERTYLFCDVDASL
DTWEKNAVHRRGLEIKPGARKCDIAELNEMREMMDLKYRSFGYGHFVLSHY
GREAGVELREDIDTVLQPGMVVSMPEMVLPEGAPGAGSYREHDLIVGEDGAEINITG
PLPGEPHNHITRN"
ORIGIN
Query Match 80.9%; Score 980; DB 6; Length 1212;
Best Local Similarity 88.0%; Pred. No. 8.2e-111;
Matches 1067; Conservative 0; Mismatches 145; Indels 0; Gaps 0;
QY 1 ATGACTGAGCAATGTTGCACTGATGAAATGGCAACAAGGCGAGAAAGATTATTCGCG 60
Db 1 ATGACTGAGCAATGTTGCACTGATGAAATGGCAACAATGATGAAAGAAATATTCGCG 60
QY 61 TTTTGGAGTCCGAGATGACCCGCGCAAAAAGAGCTTCGCGGCTGATGCGCAAGAC 120
Db 61 TTTTCCGATGCGAGATGACCGCGCCCGCAGAGTGAAGCTGCGGATGCGCAAAAC 120
QY 121 AATGTCGATGCGCGCTGTTCACTCTTATCACTGATCACTAATTCCGCTGCTG 180
Db 121 GACGTCGATGCGCGCTGTTCACTCTTATCACTGATCACTAATTCCGATTCCTG 180
QY 181 TACTGCTATTTTGGAGCGCAAGTACCGGATGCTATGACCAACAAGCCGACAGAT 240
Db 181 TACTGCTATTTTGGAGCGCAAAATACCGGATGCTATGACCAAGACCATGCGACATC 240
QY 241 TCGGCGGAGATGACGCGCGCGAGCCTGCGCGCGAGCTTGGCGCAACAATCACTAC 300
Db 241 TCGGCGGAGATGACGCGCGCGAGCCTTGGCGCGCGTACGTTGGCGCAACAATCACT 300
QY 301 ACCGACTGGCGCGCGCAATTTCTATCGCGCGCTGCGCGCACTGACCAAGGCGCAAG 360
Db 301 ACCGACTGGCGCGCGCAACTTTCTACAGGCGCTGCGCGCACTACCCCGCGCGCAAG 360
QY 361 CGCATGCGCATGAGTTTCCAGCACTGATCTGCACTTCCGCGCGCACTCCAGAGGC 420

Db 361 CGCATGCTATGAGTTGATGATCAGTGAACCTTGACTTCCGCGCGCACTCCAGAGGCG 420
QY 421 CTACCGGCGCTGAGATTGTTGATCATGACCGCAAGCCTTGATGATGATGCGCACTCAG 480
Db 421 CTGCGCGCGCTGAGATTGTTGATCATGACCGCAAGTGTGATGATGCGCACTGATGAG 480
QY 481 TCGCTGAAAGAGCAAGACTGATCCGGAAGCGCGCGGTGTGTGATCGTGGCGGCGCG 540
Db 481 TCGCTGAAAGAGCAAGACTGATCCGGAAGGTGCGGTATCTGCGACGTGCGCGGTGCG 540
QY 541 GCTTCGCGGCTGCTCATCAAGCGCGCGCTGCGCGCAAGATGAGTGGATGCGCAAC 600
Db 541 GCTTCGCTGCTGCGCGCTCATCAAGCGCGCGCTTCCGAGGACGAGGTGCGCATGCGCAAC 600
QY 601 AATGCGATGATCCGCGAATGCGCAATTCCTTCCCTTCTGAGACTGATGACACTG 660
Db 601 AATGCGATGATCCGCGAATGCGCAAGTGTTCCTTCCCTTCTGAGACTGATGACACTG 660
QY 661 ACCTGGTTCCAGTGGGATCAACACGAGCGCGCGCAATTCGGTCAACACCGCATC 720
Db 661 ACCTGGTTCCAGTGGGATCAACACGAGCGCGCGCAATTCGGTCAACACCGCATC 720
QY 721 GTGCAATCCGCGCAATCTTTCGCTCAACACTTCCCGATGATCTTCCGCTACTAC 780
Db 721 GTGCAATCCGCGCAATCTTTCGCTCAACAGCTTCCCGATGATCTTCCGCTACTAC 780
QY 781 GCGCTGAGCGCAAGCTGTTCTGCGACATGTCATGACCGCAAGCTGACATCTGAG 840
Db 781 GCGCTGAGCGCAAGCTGTTCTGCGACATGTCATGACCGCAAGCTTGTGACACTGAG 840
QY 841 AAGAACGTGGCGCTGATCCGCGCGGCTCGAGCTGATCAACGCGGCGCGCTGCAAG 900
Db 841 AAGAACGTGGCGCTGATCCGCGCGGCTCGAATCTGATCAACCGGCGCGCTGCAAG 900
QY 901 GACATCGCATGAGCTCAACAGAGTATACCGCGAGTGGACCTGCTGAAGTACCGCTCC 960
Db 901 GATATCGCATGAGCTCAACAGAGTATACCGGAGTGGAGTCTGCTGAAGTACCGCTCC 960
QY 961 TTGGCTATGCGCACTCTTTCGCGCTGTCGCACTACTAGTTCGCGGAGCGCGCTG 1020
Db 961 TTGGCTATGCGCACTCTTTCGCGCTGTCGCACTACTAGTTCGCGGAGCGCGCTG 1020
QY 1021 GAGCTGCGCGAGCAATGACACCGAGCTGAAGCGCGCGCATAGTGTCTCCATGAGCGG 1080
Db 1021 GAGCTGCGCGAGCAATGATGATCCGTGTCGACCGCGAGTGTGTCTCCATGAGCGG 1080
QY 1081 ATGATGATGCTGCGCGAGGCGATGCGCGGCTGCGCGCTATGCGAGACGACATCTG 1140
Db 1081 ATGATGATGCTGCGCGAGAGCGCTGCGCGGCTGCGCGGCTACCGAGACGACATCTG 1140
QY 1141 ATGCTGGGAGAGACGCTGCGCGAGCAATCAACCGGCTTCCGTTGGTTCGGAACAAC 1200
Db 1141 ATGCTGGGAGAGAGCGCGCAGAGCAATTAACCGATTCCCTTGGGCTGAGCAAC 1200
QY 1201 ATCATCCGCAAC 1212
Db 1201 ATCATCCGCAAC 1212
RESULT 14
AX721947 1215 bp DNA linear PAT 07-MAY-2003
LOCUS Sequence 15 from Patent EP1298213.
DEFINITION AX721947
ACCESSION AX721947
VERSION AX721947.1 GI:30422523
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE 1
AUTHORS Shao,Z., Schmuck,R., Kratzsch,P., Kenlles,J. and Weisner,H.
TITLE Variants of an erwinia-type creatinase

JOURNAL Patent: EP 1298213-A 15 02-APR-2003;
Roche Diagnostics GmbH (DE) ; F.HOFFMANN-LA ROCHE AG (CH)
FEATURES Location/Qualifiers

source
1..1215
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="variant CT2m28"
1..1215
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAD90220.1"
/db_xref="GI:30422524"
/translation="MTDMLHVMKMGKEKVSPESDAEMTRSDVRRMAENDVDA
ALFTSYHCINITYSGFLICYFGKRYGAVITDQDAHTTISAGIDGQPRRSFGNITTTD
WRNDYQAVRQLTPGARRIIGIEFHVLDLPERLTLEALPGVEFVIGPSPMMRTVK
SLEBQCLIREGARI CDVGGA CVAAVKAGVPEHEVA IATTNVAREIAGSFPVELMD
TWTWFGSINTDGAHPVTRNIIVOSGDILSLNTPMI FGYYTALERTLFCDHVDASL
DTWEKNVAVHRRELILIKPGARKDIALENEMYREMDLTKRSPFGYHSFGVLSHY
GRBAGVELLEDIDTVLQPGMVYSMEPMWMLPEGA PGAGYREHDLIIVGEDGMENTG
PFGPEPHNITRN"

ORIGIN

Query Match 80.9%; Score 980; DB 6; Length 1215;
Best Local Similarity 88.0%; Pred. No. 8.2e-111;
Matches 1067; Conservative 0; Mismatches 145; Indels 0; Gaps 0;

QY 1 ATGACTGACGACATGTTGCACTGATGAAATGGCACACGCGCAGAAAGATTATTCGCG 60
DB 1 ATGACTGACGACATGTTGCACTGATGAAATGGCACACATGTTGAGAAATATTCGCG 60
QY 61 TTTTGGGATGCGGAGATGACCGCGCGCAAAAGAGTTGCGCGCTGATGCGCAAGAC 120
DB 61 TTTTGGGATGCGGAGATGACCGCGCGCAAAAGAGTTGCGCGCTGATGCGCAAGAC 120
QY 61 TTTTGGGATGCGGAGATGACCGCGCGCAAAAGAGTTGCGCGCTGATGCGCAAGAC 120
DB 61 TTTTGGGATGCGGAGATGACCGCGCGCAAAAGAGTTGCGCGCTGATGCGCAAGAC 120
QY 121 AATGTCATGCGCGCGCTGTTCACTCTTATCACTGATCACTACTATTCGCGCTGCTG 180
DB 121 GACGTCGACGCTGCGCTGTTCACTCTTATCACTGATCACTACTATTCGCGCTGCTG 180
QY 181 TACTGCTATTTGCGACGCAAGTACGCGATGTCATGACACCAACGCGCAAGCAT 240
DB 181 TACTGCTATTTGCGACGCAAGTACGCGATGTCATGACACCAACGCGCAAGCAT 240
QY 181 TACTGCTATTTGCGACGCAAGTACGCGATGTCATGACACCAACGCGCAAGCAT 240
DB 181 TACTGCTATTTGCGACGCAAGTACGCGATGTCATGACACCAACGCGCAAGCAT 240
QY 241 TGGGCGGATGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
DB 241 TGGGCGGATGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
QY 241 TGGGCGGATGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
DB 241 TGGGCGGATGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300
QY 301 ACCGACTGCGCGCGCGCAATTTCTATCGCGCGCGCGCGCGCGCGCGCGCGCG 360
DB 301 ACCGACTGCGCGCGCGCAATTTCTATCGCGCGCGCGCGCGCGCGCGCGCGCG 360
QY 361 CGCATGCGATGAGTTGCAACGCTCAATCTGCACTTCGCGCGCGCGCGCGCGCG 420
DB 361 CGCATGCGATGAGTTGCAACGCTCAATCTGCACTTCGCGCGCGCGCGCGCGCG 420
QY 361 CGCATGCGATGAGTTGCAACGCTCAATCTGCACTTCGCGCGCGCGCGCGCGCG 420
DB 361 CGCATGCGATGAGTTGCAACGCTCAATCTGCACTTCGCGCGCGCGCGCGCGCG 420
QY 421 CTACCGCGCGCTGAGTTGCTGCAATGACCGCGCGCGCGCGCGCGCGCGCGCG 480
DB 421 CTACCGCGCGCTGAGTTGCTGCAATGACCGCGCGCGCGCGCGCGCGCGCGCG 480
QY 421 CTACCGCGCGCTGAGTTGCTGCAATGACCGCGCGCGCGCGCGCGCGCGCGCG 480
DB 421 CTACCGCGCGCTGAGTTGCTGCAATGACCGCGCGCGCGCGCGCGCGCGCGCG 480
QY 481 TGGCTCGAAGACGAAAGTGAATCGCGAAGGCGCGCGCGCGCGCGCGCGCGCG 540
DB 481 TGGCTCGAAGACGAAAGTGAATCGCGAAGGCGCGCGCGCGCGCGCGCGCGCG 540
QY 541 GCTGCGCGCTGCGCATCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
DB 541 GCTGCGCGCTGCGCATCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
QY 601 AATGCGATGATCGCGAGATCGCGCAATGCTTCCCTTCTGCGAGCTGAAGCACTGG 660
DB 601 AATGCGATGATCGCGAGATCGCGCAATGCTTCCCTTCTGCGAGCTGAAGCACTGG 660
QY 661 AACTGCTTCAATCGCGGATCAACACCGAGCGCGCGCGCGCGCGCGCGCGCG 720
DB 661 AACTGCTTCAATCGCGGATCAACACCGAGCGCGCGCGCGCGCGCGCGCGCG 720

DB 661 AACTGCTTCAATCGCGGATCAACACCGAGCGCGCGCGCGCGCGCGCGCGCG 720
QY 721 GTGCAATCGCGGACATCTCTTGTGCAACCTTCCGATGATCTTGGCTCTACACC 780
DB 721 GTGCAATCGCGGACATCTCTTGTGCAACCTTCCGATGATCTTGGCTCTACACC 780
QY 781 GGGCTGAGACGACGCTGTTCTGCGACGATGTCATGATGAGCGCGCTCGACATCTGGAG 840
DB 781 GGGCTGAGACGACGCTGTTCTGCGACGATGTCATGATGAGCGCGCTCGACATCTGGAG 840
QY 841 AAGAACTGCGCGTGTGATGCGCGCGCGCTGAGCTGATCAAGCGCGCGCGCTGCAAG 900
DB 841 AAGAACTGCGCGTGTGATGCGCGCGCGCTGAGCTGATCAAGCGCGCGCGCTGCAAG 900
QY 901 GACATGCGCATGAGCTGCAACGAGATGTCACCGGATGAGGACCTGCTGAAGTACCGCTCC 960
DB 901 GACATGCGCATGAGCTGCAACGAGATGTCACCGGATGAGGACCTGCTGAAGTACCGCTCC 960
QY 961 TTCGCTATGAGCGACCTCTTCCGCGCGCTGTGTCACCTACTACGATGCGAGCGCGCG 1020
DB 961 TTCGCTATGAGCGACCTCTTCCGCGCGCTGTGTCACCTACTACGATGCGAGCGCGCG 1020
QY 1021 GAGCTGCGCGAGACATGCAACCGAGCTGAAGCCCGGATGAGTGTCTCATGAGCGG 1080
DB 1021 GAGCTGCGCGAGACATGCAACCGAGCTGAAGCCCGGATGAGTGTCTCATGAGCGG 1080
QY 1081 ATGCTGATGCTGCGCGAGGACATGCGCGCGCTGCGCGCTGATGCGACGACATCTCG 1140
DB 1081 ATGCTGATGCTGCGCGAGGACATGCGCGCGCTGCGCGCTGATGCGACGACATCTCG 1140
QY 1141 ATGCTGATGCTGCGCGAGGACATGCGCGCGCTGCGCGCTGATGCGAGCAAC 1200
DB 1141 ATGCTGATGCTGCGCGAGGACATGCGCGCGCTGCGCGCTGATGCGAGCAAC 1200
QY 1201 ATCATCCGCAAC 1212
DB 1201 ATCATCCGCAAC 1212

RESULT 15
AX721953
LOCUS AX721953 1212 bp DNA linear PAT 07-MAY-2003
DEFINITION Sequence 21 from Patent EP1298213.
ACCESSION AX721953
VERSION AX721953.1 GI:30422529
KEYWORDS
SOURCE
ORGANISM
synthetic construct
other sequences; artificial sequences.
REFERENCES
1 Shao, Z., Schumack, R., Kratzsch, P., Kenkies, J. and Weisner, H.
AUTHORS
TITLE
Variants of an erwinia-type creatinase
JOURNAL
Patent: EP 1298213-A 21 02-APR-2003;
Roche Diagnostics GmbH (DE) ; F.HOFFMANN-LA ROCHE AG (CH)
FEATURES Location/Qualifiers

source
1..1212
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="variant CT2d7"
1..1212
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAD90223.1"
/db_xref="GI:30422530"
/translation="MTDMLHVMKMGKEKVSPESDAEMTRSDVRRMAENDVDA
ALFTSYHCINITYSGFLICYFGKRYGAVITDQDAHTTISAGIDGQPRRSFGNITTTD
WRNDYQAVRQLTPGARRIIGIEFHVLDLPERLTLEALPGVEFVIGPSPMMRTVK
SLEBQCLIREGARI CDVGGA CVAAVKAGVPEHEVA IATTNVAREIAGSFPVELMD
TWTWFGSINTDGAHPVTRNIIVOSGDILSLNTPMI FGYYTALERTLFCDHVDASL
DTWEKNVAVHRRELILIKPGARKDIALENEMYREMDLTKRSPFGYHSFGVLSHY

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 7, 2005, 19:27:19 ; Search time 728 Seconds
(without alignments)
9855.394 Million cell updates/sec

Title: US-10-807-228A-2

Perfect score: 1212
Sequence: 1 ATACTGACGACATGTGCA.....AACACAACATCATCCGAAC 1212

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1210.4	99.9	1212	2	AAT61367 Creatinin
2	1210.4	99.9	1212	2	AAT38807 Creatine
3	1210.4	99.9	1212	2	AAV5569 Strable cr
4	1210.4	99.9	1212	4	AA514742 Creatine
5	1210.4	99.9	1212	6	AA199856 Alkaligen
6	1210.4	99.9	1212	6	ABA93696 Creatine
7	1188	98.0	1215	2	AAT13291 Creatine
8	984.8	81.3	1215	8	ACC69514 Erwila s
9	984.8	81.3	1215	8	ACC69519 Mutant Er
10	981.6	81.0	1215	8	ACC69515 Mutant Er
11	980	80.9	1212	8	ACC69517 Mutant Er
12	980	80.9	1212	8	ACC69516 Mutant Er
13	980	80.9	1215	8	ACC69518 Mutant Er
14	978.4	80.7	1212	8	ACC69520 Mutant Er
15	976.8	80.6	1212	8	ACC69520 Mutant Er
16	643.4	53.1	1209	2	AAT05502 Thermosca
17	638.8	52.7	1212	1	AAAB0274 Sequence
18	637.2	52.6	1212	1	AAAB1271 Sequence
19	564.6	46.6	1134	1	AAAT0877 Sequence
20	459.4	37.9	1233	1	AAAN90635 Sequence

ALIGNMENTS

RESULT 1	
1	AAT61367 standard; DNA; 1212 BP.
2	AAV66090
3	ABD05567
4	ABD05831
5	ABD05710
6	ABD05710
7	ADG73350
8	ADG73352
9	ABQ44816
10	ABQ44817
11	AAV08896
12	ABE1294
13	AAAD17184
14	AAAD17186
15	ACA23861
16	AAA05948
17	ACA36793
18	ABR03041.18
19	ADG91713
20	ADG91695
21	ADN40876
22	ACC44573
23	AAZ06824
24	AAH74537
25	AAH74537
26	AAH74537
27	AAH74537
28	AAH74537
29	AAH74537
30	AAH74537
31	AAH74537
32	AAH74537
33	AAH74537
34	AAH74537
35	AAH74537
36	AAH74537
37	AAH74537
38	AAH74537
39	AAH74537
40	AAH74537
41	AAH74537
42	AAH74537
43	AAH74537
44	AAH74537
45	AAH74537

21	319.2	26.3	1282	2	AAV66090
22	74.8	6.2	858	11	ABD05567
23	74.8	6.2	963	11	ABD05831
24	74.8	6.2	1176	11	ABD05710
25	73.2	6.0	786	10	ADG73350
26	73.2	6.0	786	10	ADG73352
27	68	5.6	985	6	ABQ44816
28	68	5.6	985	6	ABQ44817
29	68	5.6	3157	2	AAV08896
30	67.8	5.6	1227	6	ABE1294
31	67.2	5.5	65140	4	AAAD17184
32	67.2	5.5	125401	4	AAAD17186
33	66.6	5.5	2052	8	ACA23861
34	65.2	5.4	690	3	AAA05948
35	65.2	5.4	2712	8	ACA36793
36	65.2	5.4	110000	6	ABR03041.18
37	65	5.4	2253	13	ADG91713
38	64.6	5.3	164051	13	ADG91695
39	64.6	5.3	9546	12	ADN40876
40	64.4	5.3	2037	8	ACC44573
41	64	5.3	1294	2	AAZ06824
42	64	5.3	1294	4	AAH74537
43	63.6	5.2	690	6	AAH74537
44	63.4	5.2	1161	8	ACA23536
45	63.2	5.2	1284	8	ADA70934

21	319.2	26.3	1282	2	AAV66090
22	74.8	6.2	858	11	ABD05567
23	74.8	6.2	963	11	ABD05831
24	74.8	6.2	1176	11	ABD05710
25	73.2	6.0	786	10	ADG73350
26	73.2	6.0	786	10	ADG73352
27	68	5.6	985	6	ABQ44816
28	68	5.6	985	6	ABQ44817
29	68	5.6	3157	2	AAV08896
30	67.8	5.6	1227	6	ABE1294
31	67.2	5.5	65140	4	AAAD17184
32	67.2	5.5	125401	4	AAAD17186
33	66.6	5.5	2052	8	ACA23861
34	65.2	5.4	690	3	AAA05948
35	65.2	5.4	2712	8	ACA36793
36	65.2	5.4	110000	6	ABR03041.18
37	65	5.4	2253	13	ADG91713
38	64.6	5.3	164051	13	ADG91695
39	64.6	5.3	9546	12	ADN40876
40	64.4	5.3	2037	8	ACC44573
41	64	5.3	1294	2	AAZ06824
42	64	5.3	1294	4	AAH74537
43	63.6	5.2	690	6	AAH74537
44	63.4	5.2	1161	8	ACA23536
45	63.2	5.2	1284	8	ADA70934

21	319.2	26.3	1282	2	AAV66090
22	74.8	6.2	858	11	ABD05567
23	74.8	6.2	963	11	ABD05831
24	74.8	6.2	1176	11	ABD05710
25	73.2	6.0	786	10	ADG73350
26	73.2	6.0	786	10	ADG73352
27	68	5.6	985	6	ABQ44816
28	68	5.6	985	6	ABQ44817
29	68	5.6	3157	2	AAV08896
30	67.8	5.6	1227	6	ABE1294
31	67.2	5.5	65140	4	AAAD17184
32	67.2	5.5	125401	4	AAAD17186
33	66.6	5.5	2052	8	ACA23861
34	65.2	5.4	690	3	AAA05948
35	65.2	5.4	2712	8	ACA36793
36	65.2	5.4	110000	6	ABR03041.18
37	65	5.4	2253	13	ADG91713
38	64.6	5.3	164051	13	ADG91695
39	64.6	5.3	9546	12	ADN40876
40	64.4	5.3	2037	8	ACC44573
41	64	5.3	1294	2	AAZ06824
42	64	5.3	1294	4	AAH74537
43	63.6	5.2	690	6	AAH74537
44	63.4	5.2	1161	8	ACA23536
45	63.2	5.2	1284	8	ADA70934

21	319.2	26.3	1282	2	AAV66090
22	74.8	6.2	858	11	ABD05567
23	74.8	6.2	963	11	ABD05831
24	74.8	6.2	1176	11	ABD05710
25	73.2	6.0	786	10	ADG73350
26	73.2	6.0	786	10	ADG73352
27	68	5.6	985	6	ABQ44816
28	68	5.6	985	6	ABQ44817
29	68	5.6	3157	2	AAV08896
30	67.8	5.6	1227	6	ABE1294
31	67.2	5.5	65140	4	AAAD17184
32	67.2	5.5	125401	4	AAAD17186
33	66.6	5.5	2052	8	ACA23861
34	65.2	5.4	690	3	AAA05948
35	65.2	5.4	2712	8	ACA36793
36	65.2	5.4	110000	6	ABR03041.18
37	65	5.4	2253	13	ADG91713
38	64.6	5.3	164051	13	ADG91695
39	64.6	5.3	9546	12	ADN40876
40	64.4	5.3	2037	8	ACC44573
41	64	5.3	1294	2	AAZ06824
42	64	5.3	1294	4	AAH74537
43	63.6	5.2	690	6	AAH74537
44	63.4	5.2	1161	8	ACA23536
45	63.2	5.2	1284	8	ADA70934

21	319.2	26.3	1282	2	AAV66090
22	74.8	6.2	858	11	ABD05567
23	74.8	6.2	963	11	ABD05831
24	74.8	6.2	1176	11	ABD05710
25	73.2	6.0	786	10	ADG73350
26	73.2	6.0	786	10	ADG73352
27	68	5.6	985	6	ABQ44816
28	68	5.6	985	6	ABQ44817
29	68	5.6	3157	2	AAV08896
30	67.8	5.6	1227	6	ABE1294
31	67.2	5.5	65140	4	AAAD17184
32	67.2	5.5	125401	4	AAAD17186
33	66.6	5.5	2052	8	ACA23861
34	65.2	5.4	690	3	AAA05948
35	65.2	5.4	2712	8	ACA36793
36	65.2	5.4	110000	6	ABR03041.18
37	65	5.4	2253	13	ADG91713
38	64.6	5.3	164051	13	ADG91695
39	64.6	5.3	9546	12	ADN40876
40	64.4	5.3	2037	8	ACC44573
41	64	5.3	1294	2	AAZ06824
42	64	5.3	1294	4	AAH74537
43	63.6	5.2	690	6	AAH74537
44	63.4	5.2	1161	8	ACA23536
45	63.2	5.2	1284	8	ADA70934

21	319.2	26.3	1282	2	AAV66090
22	74.8	6.2	858	11	ABD05567
23	74.8	6.2	963	11	ABD05831
24	74.8	6.2	1176	11	ABD05710
25	73.2	6.0	786	10	ADG73350
26	73.2	6.0	786	10	ADG73352
27	68	5.6	985	6	ABQ44816
28	68	5.6	985	6	ABQ44817
29	68	5.6	3157	2	AAV08896
30	67.8	5.6	1227	6	ABE1294
31	67.2	5.5	65140	4	AAAD17184
32	67.2	5.5	125401	4	AAAD17186
33	66.6	5.5	2052	8	ACA23861
34	65.2	5.4	690	3	AAA05948
35	65.2	5.4	2712	8	ACA36793
36	65.2	5.4	110000	6	ABR03041.18
37	65	5.4	2253	13	ADG91713
38	64.6	5.3	164051	13	ADG91695
39	64.6	5.3	9546	12	ADN40876
40	64.4	5.3	2037	8	ACC44573
41	64	5.3	1294	2	AAZ06824
42	64	5.3	1294	4	AAH74537
43	63.6	5.2	690	6	AAH74537
44	63.4	5.2	1161	8	ACA23536
45	63.2	5.2	1284	8	ADA70934

```

xx Sequence 1212 BP; 244 A; 400 C; 358 G; 210 T; 0 U; 0 Other;
SQ Query Match 99.9%; Score 1210.4; DB 2; Length 1212;
Best Local Similarity 99.9%; Pred. No. 4,2e-196;
Matches 1211; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGACGTGACGATGTTGACGATGTAATGACCAAGGCGAGAAAGATTATTCGCCG 60
DB 1 ATGACGTGACGATGTTGACGATGTAATGACCAAGGCGAGAAAGATTATTCGCCG 60
QY 61 TTTTGGATGCGGAGATGACCCGCGCCAAAAGAGTTGCGGCTGATGAGCAAGAC 120
DB 61 TTTTGGATGCGGAGATGACCCGCGCCAAAAGAGTTGCGGCTGATGAGCAAGAC 120
QY 121 AATGTGATGCGGCGCTGTTCACTCTTATATCACTGATCACTACTATTCCGGCTGGCTG 180
DB 121 AATGTGATGCGGCGCTGTTCACTCTTATATCACTGATCACTACTATTCCGGCTGGCTG 180
QY 181 TACGTGCTATTTGGAAGCGAAGTACGGCATGATGACACAAAGCCGACGAGATT 240
DB 181 TACGTGCTATTTGGAAGCGAAGTACGGCATGATGACACAAAGCCGACGAGATT 240
QY 241 TCGGCGGCGATGACGCGCGCGAGCCCTGCGCGCGGAGCTTCCGCGCAATCACTAC 300
DB 241 TCGGCGGCGATGACGCGCGCGAGCCCTGCGCGCGGAGCTTCCGCGCAATCACTAC 300
QY 301 ACCGACTGCGCGCGCGCAATTTCTATCGCGCGCTGCGCGCTGACGACGAGCGCCAG 360
DB 301 ACCGACTGCGCGCGCGCAATTTCTATCGCGCGCTGCGCGCTGACGACGAGCGCCAG 360
QY 361 CGCATGCGGATGAGTTGACCAAGTCATCTGACCTTCGCGCGCTGAGAGAAAGCC 420
DB 361 CGCATGCGGATGAGTTGACCAAGTCATCTGACCTTCGCGCGCTGAGAGAAAGCC 420
QY 421 CTACCGGCGGCTGAGTTGCTGCAATCAAGCCAGCCCTGATGATGATGCGACATCAAG 480
DB 421 CTACCGGCGGCTGAGTTGCTGCAATCAAGCCAGCCCTGATGATGATGCGACATCAAG 480
QY 481 TCGCTCGAAGAGCAGAAAGCTGATCCGCGAAGCGCGCGCTGATGATGATGATGATG 540
DB 481 TCGCTCGAAGAGCAGAAAGCTGATCCGCGAAGCGCGCGCTGATGATGATGATGATG 540
QY 541 GCGTCCGCGGCTGCGCATCAAGCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
DB 541 GCGTCCGCGGCTGCGCATCAAGCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
QY 601 AATGCGATGATCCGCGAGATCCGCAAAATGTTCCCTTCTGATGATGATGATGATG 660
DB 601 AATGCGATGATCCGCGAGATCCGCAAAATGTTCCCTTCTGATGATGATGATGATG 660
QY 661 ACTGTGTTCACTGCGGCGATCAACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCATC 720
DB 661 ACTGTGTTCACTGCGGCGATCAACCGAGCGCGCGCGCGCGCGCGCGCGCGCGCATC 720
QY 721 GTGCAATCCGCGGAGATCCCTTCTGCAACCTTCCGATATCTTCCGCTATCAACCC 780
DB 721 GTGCAATCCGCGGAGATCCCTTCTGCAACCTTCCGATATCTTCCGCTATCAACCC 780
QY 781 GCGCTGAGACGACCGCTGTTCTGCAACATGATGATGATGATGATGATGATGATG 840
DB 781 GCGCTGAGACGACCGCTGTTCTGCAACATGATGATGATGATGATGATGATGATG 840
QY 841 AAGAAAGTGGCGGTGATGCGCGCGGCTGAGCTGATCAAGCCGCGCGCGCTGCAAG 900
DB 841 AAGAAAGTGGCGGTGATGCGCGCGGCTGAGCTGATCAAGCCGCGCGCGCTGCAAG 900
QY 901 GACATGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
DB 901 GACATGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 960
QY 961 TTCGCTATGCGCATCTCTTCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
DB 961 TTCGCTATGCGCATCTCTTCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020

```

```

DB 961 TTCGCTATGCGCATCTCTTCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
QY 1021 GAGCTGCCCGAGACATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
DB 1021 GAGCTGCCCGAGACATGATGATGATGATGATGATGATGATGATGATGATGATG 1080
QY 1081 ATGTGATGCTGCGGAGGATGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 1140
DB 1081 ATGTGATGCTGCGGAGGATGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 1140
QY 1141 ATGCTGCGGAGAGACGCTGCGGAGACATCAACCGGCTTCCGCTGCTGCGGAGAC 1200
DB 1141 ATGCTGCGGAGAGACGCTGCGGAGACATCAACCGGCTTCCGCTGCTGCGGAGAC 1200
QY 1201 ATCATCCGCAAC 1212
DB 1201 ATCATCCGCAAC 1212

RESULT 2
AAT38807
ID AAT38807 standard; DNA; 1212 BP.
XX
AC AAT38807;
XX
DT 17-OCT-2003 (revised)
DT 02-MAR-1998 (first entry)
XX
DE Creatine amidinohydrolyase gene.
XX
KM Creatine amidinohydrolyase enzyme; sarcosine; urea; dye; Km; ds.
OS Alcaligenes faecalis; - strain TE3581 (FERM P-14237).
XX
FH Key Location/Qualifiers
FT CDS 1..1212
FT /tag= a
FT /product= "Creatine amidinohydrolyase"
FT /transl_except= (pos:433..435,aa:Glu)
FT /note= "no stop codon given"
XX
EP790303-A1.
XX
PD 20-AUG-1997.
XX
PF 13-FEB-1997; 97EP-00102270.
XX
PR 13-FEB-1996; 96JP-00025435.
XX
PA (TOYM ) TOYO BOSEKI KK.
XX
PI Sogabe A, Hattori T, Nishiya Y, Kawamura Y;
DR WPI; 1997-404731/38.
DR P-PsDB; AAM22893.
XX
PT Creatine amidinohydrolyase enzyme with low Km - for use in assay for
PT creatine.
XX
PS Disclosure; Page 14-15; 21pp; English.
XX
CC A novel creatine amidinohydrolyase enzyme has been developed which
CC catalyses the reaction of creatine with water to form sarcosine and urea,
CC is stable at temperatures of up to 50 degrees Celsius (pH 7.5, 30
CC minutes) and has an optimum temperature of 40-50 degrees Celsius, optimum
CC pH of 8-9, a Km value for creatine of 3.5-10 mM in a binding assay using
CC sarcosine oxidase and peroxidase, a molecular weight of 43 kD by SDS-PAGE
CC and an isoelectric point of 3.5. The present sequence encodes creatine
CC amidinohydrolyase derived from Alcaligenes faecalis strain TE3581 (FERM P-
CC 14237), which is the wild type creatine amidinohydrolyase to be mutated in
CC the present invention. The enzyme can be used to determine creatine in a
CC sample by measuring the absorbance of a dye formed by reacting the sample
CC with a reagent, comprising the enzyme, sarcosine oxidase and a

```